

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 18:14:11 ; Search time 2037.42 Seconds

(without alignments)  
12546.865 Million cell updates/sec

Title: US-09-841-963A-1

Perfect score: 1894

Sequence: 1 gctgactctctccacgac.....ataagatactagagactg 1894

Scoring table: IDENTITY\_NUC  
Gap 10.0 , Gapext 1.0

Searched: 13736207 segs, 6748477542 residues 27472414

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlun:\*  
5: em\_estlov:\*  
6: em\_estcpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gD\_estl:\*  
10: gD\_estl2:\*  
11: gD\_hic:\*  
12: gD\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_lun:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vit:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	781.4	41.3	799	10	BT770453 603055418
2	712.4	37.6	730	10	BT772270 603055851
3	682.6	36.0	991	10	BT779568 602328182
4	659.6	34.8	721	9	AT745526
5	617.4	32.6	640	10	BT589055 nacc29604
6	604.4	31.9	843	10	BT178499
7	596.6	31.5	927	10	BT476482
8	590	31.2	933	10	BT563288
9	584.8	30.9	945	10	BT871340
10	576.4	30.4	625	10	BT446190
11	573	30.3	950	10	BT006343
12	549.6	29.0	899	10	BT328411
13	549.6	29.0	606	9	AA662164
14	525.8	27.8	652	9	AA662164
15	504.4	26.7	514	9	AA662204
16	467.6	24.7	809	10	BT567992
17	465	24.6	601	9	AA971121

c	18	463.8	24.5	512	9	AT668916
	19	463.8	24.5	737	10	BT077619
	20	460.6	24.3	539	10	BT57615
	21	448.6	23.7	591	10	BT085247
	22	431.8	22.8	458	9	AT991223
	23	417.4	22.0	430	9	AT979170
	24	396.2	20.9	418	9	AA298436
	25	381	20.1	943	10	BT676137
	26	372.2	19.7	392	9	AA639821
	27	352.6	18.6	361	9	AA513942
	28	349.8	18.5	353	10	BT003063
	29	346.8	18.3	381	9	AT859010
	30	345.6	18.2	386	9	AT685592
	31	345.6	18.2	548	9	AA973391
	32	344	18.2	366	9	AT687669
	33	341.6	18.0	2004	10	BT579116
	34	323.6	17.1	366	9	AA662521
	35	316	16.7	335	9	AA513274
	36	306.8	16.2	859	9	AA548598
	37	302.6	15.7	628	10	BT072997
	38	297.4	15.7	369	9	AA228304
	39	289.6	15.3	784	10	BT072422
	40	278.6	14.7	441	9	AT686414
	41	275.4	14.5	504	9	AA930371
	42	256.6	13.5	952	10	BT539428
	43	236.6	12.5	348	9	AT922279
	44	235	12.4	875	10	BT56737
	45	231	12.2			

## ALIGNMENTS

RESULT 1  
BT770453  
LOCUS  
DEFINITION  
BT770453  
603055418F1 NIH\_MGC\_122 Homo sapiens CDNA clone IMAGE:5204692 5',  
mRNA sequence.

ACCESSION  
BT770453  
VERSION  
BT770453.1  
KEYWORDS  
EST,  
SOURCE  
human.  
ORGANISM  
Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L1AM1513 row: d column: 05  
High quality sequence stop: 799.  
Location/Qualifiers  
1..799  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5204692"  
/clone\_1lb="NIH\_MGC\_122"  
/lab\_host="DH10B"  
/note="Organ: pooled lung and spleen; Vector: PCMV-SPORT6;  
site\_1: NotI; site\_2: EcoRV (destroyed); RNA source  
anonymous pool of 24 week female lung, 16 week female  
spleen, and 20-22 week male spleens. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.4 kb, insert size

## FEATURES

source

range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library.  
 BASE COUNT 135 a 297 c 248 g 119 t  
 ORIGIN

Query Match 41.3%; Score 781.4; DB 10; Length 799;  
 Best Local Similarity 98.6%; Pred. No. 5.7e-135;  
 Matches 788; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1 gctgactctctccagacatctctcactctcgtctgctccacactgccccacagacc 60  
 1 GTGTAGCTCTCTCCACACATTCCTGCGCTGCTCCACACTGCCCCACAGACCC 60  
 Db 1 gctgactctctccagacatctctcactctcgtctgctccacactgccccacagacc 60  
 61 agtctcacaagctctcagctccctcagagccctcagatttggtgctctgacagtg 120  
 61 AGTCTCTCAAGCCCTGCTGCGCTGCTCCACAGCCCTCAGTTGGCTTGGCCACGGTG 120  
 Db 61 agtctcacaagctctcagctccctcagagccctcagatttggtgctctgacagtg 120  
 121 ccagcagagcagccctctggtctggttagggatccctcagcagcagcagccctgagac 180  
 121 CCACAGAGCAGCCCTGCGCTGCGGTAGGGACTCCCTACAGCGACGCGCCCTGAGACC 180  
 Db 121 ccagcagagcagccctctggtctggttagggatccctcagcagcagcagccctgagac 180  
 181 gccacacagcagccctctgaggtgtgagcagcccccagctgagccactgagtgctctc 240  
 181 TCAGAGGGCCACCCCTTGAGAGGTGGCCAGGCCCCAGTGGCCAGCTGAGTGCCTCT 240  
 Db 181 gccacacagcagccctctgaggtgtgagcagcccccagctgagccactgagtgctctc 240  
 241 gccacacagcagccctctgaggtgtgagcagcccccagctgagccactgagtgagcagc 300  
 241 GCCACACAGCCCTGCTGAGCCCTGTTGCTGCTGACCCCCAGATGCTGCTGAGACAGC 300  
 Db 241 gccacacagcagccctctgaggtgtgagcagcccccagctgagccactgagtgagcagc 300  
 301 cagtgagcagcagctctcagcagcagcctctccagcagcagcctgagtgagtgagcag 360  
 301 CAGTGGCCTCAGCTGCCACACCTCTCCCGCCCTCAAGTTGGCACTGCACGACAGACAG 360  
 Db 301 cagtgagcagcagctctcagcagcagcctctccagcagcagcctgagtgagtgagcag 360  
 361 ctcccttgaggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 420  
 361 CTCCTTGAGGACACAGGAGCTAACACACAGCCGCGACCCCAACACAGCAGCGCATGGG 420  
 Db 361 ctcccttgaggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 420  
 421 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480  
 421 CAGCGCCAGCCCGGCTGTGAGCAGCTATCCCGCAGCCACTCTGCTGCCCCCGACAC 480  
 Db 421 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480  
 481 ggtctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 540  
 481 GGTCTCGCGGACACAGCTTGGACAAAGCGCGCACCGGGGAGTGGCTCTCGAGACAGGGA 540  
 Db 481 ggtctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 540  
 541 ctgagagtcacagtcacagtcacagtcacagtcacagtcacagtcacagtcacagtcac 600  
 541 CTGAGAGTCCAGTCCACACCCCGCACAGCCAGAGGGCCCTGTCCTCTACCTCTCTCTA 600  
 Db 541 ctgagagtcacagtcacagtcacagtcacagtcacagtcacagtcacagtcacagtcac 600  
 601 ctctgacatgctgtaccccttgagagcagcagcctgagcagcagcagcagcagcagcag 660  
 601 CTTTGACATGCTGTACCTTGAGACAGCAGCTGGGAGGCAAGGCCCCCTGGGGCCACAG 660  
 Db 601 ctctgacatgctgtaccccttgagagcagcagcctgagcagcagcagcagcagcagcag 660  
 661 tcggagagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
 661 TCGGAGAGAGCAGCTTACAGAGCCTTACAGAGTCCCGGTCATTGACACCAAGCCCAAC 720  
 Db 661 tcggagagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
 721 gggcagccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780  
 721 GGGCAGCCTTGAGCTTGGTCCCGGGCGGTGACCTTGGAGAGCAGCTGCTGGAGCAGGT 780  
 Db 721 gggcagccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780  
 781 gcaatccatggt 799  
 781 GCAATCCATGGT 799  
 Db 781 gcaatccatggt 799

ACCESSION mRNA sequence.  
 B1772270  
 VERSION B1772270.1 GI:15763848  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC http://mgc.ncl.nih.gov/  
 1 (bases 1 to 730)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L14M1514 row: f column: 13  
 High quality sequence stop: 729.  
 Location/Qualifiers  
 1. 730  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5205132"  
 /clone\_11b="NIH\_MGC\_122"  
 /lab\_host="DH10B"  
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library."

BASE COUNT 124 a 265 c 237 g 104 t  
 ORIGIN

Query Match 37.6%; Score 712.4; DB 10; Length 730;  
 Best Local Similarity 98.5%; Pred. No. 3.3e-122;  
 Matches 719; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

68 caagcctgtgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 127  
 1 CAAGCCTGTGCTGCAAGCTCCCTGCAAGCCCTGCAAGCTTGGGCTTGCCACGCTCCAG 60  
 Db 68 caagcctgtgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 127  
 128 gcaagcctgtgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 187  
 61 GCAAGCCTGTGCTGCAAGCTCCCTGCAAGCCCTGCAAGCTTGGGCTTGCCACGCTCCAG 120  
 Db 128 gcaagcctgtgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 187  
 188 gccaccccttgaggtgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 247  
 188 GCCACCCCTTGAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
 Db 188 gccaccccttgaggtgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 247  
 248 gccctgtgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 307  
 181 GCCCTGTGCTGCAAGCTCCCTGCAAGCCCTGCAAGCTTGGGCTTGCCACGCTCCAG 180  
 Db 248 gccctgtgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 307  
 308 ctgagctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 367  
 241 CTCAGCTGCCACACCTCTTCCGGGCTTGAGAGTGGCATGACAGCAGCAGCTCCAG 300  
 Db 308 ctgagctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 367  
 368 ggcacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 427  
 301 GGCACACAGCAGCTTAACAGACAGCGCGCAGCCCAACAGCAGCAGCGCATGGGACAGGCC 360  
 Db 368 ggcacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 427  
 428 agcccggtctgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 487

Db	361	AGCCGGGGCTGAGCAGGGTATCCCCCAGCACCTTCGTGTCGCCCCCGACAGGGTGTGC	420
QY	488	cggacagccttggagaagcgcgcagcgggggcagltcggtctcgagaagacggacttgaagt	547
Db	421	CGGACAGGCTTGGAGAAAGCGCGCACCGGGGGGACGTGCTTCGAAGACAGGGAGACTGGAGT	480
QY	548	cccaatgcacccgcgcagcccgagagagggcctgtccgcttcgacttcacttcactttac	607
Db	461	CCCAATCCACCCGCCGACGCCGACGAGGGCCTGTCCGCTTTCACCTCTACTTTGAC	540
QY	608	atgtctgtaccccttgagagcagcagctctgtggcagccaaagccctctgtggcagcagtcgtggag	667
Db	541	ATGCTGTACCCCTGAGAGACAGCAGCTGGGCGACGCCCTCGGGGGCCAGCAGTCTGGAG	600
QY	668	gagcccaacttgacaaaccttgagcagtgcccggtcatcttgacagccaaagcccccagcgggcacc	727
Db	601	GAGCCACCTCGAGAGACCTCTGACGATGCCCCGGTCATTGACAGCCAAAGCCCCAACGGGGCAGC	660
QY	728	cttgcagcttgcgtcccccggcggtgcagcacttgagagagcaatcgcttgagcaggtgcagtcc	787
Db	661	CTGGACCTTGGTCCCGGGGGGCTGACCTTTGAGAGAGACACTGCTGTGAGCAGGTGCAGTCC	720
QY	788	atgctgtgtcgtg 797	
Db	721	ATGCTGTGCTGG 730	

RESULT	3
BGI79568	
LOCUS	
DEFINITION	BGI79568 991 bp mRNA linear EST 06-FEB-2001
ACCESSION	G03328182P1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4429413 5',
VERSION	mRNA sequence.
KEYWORDS	BGI79568 BGI79568.1 GI:12686271
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	1 (bases 1 to 991)
COMMENT	NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: DCTD/DP cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLM10181 row: d column: 22 High quality sequence stop: 655.

FEATURES	SOURCE	Location/Qualifiers
		1..991
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone_image="4428413"
		/clone_11b="NH_MGC_91"
		/tissue_type="adenocarcinoma, cell line"
		/lab_host="DH10B (phage-resistant)"
		/note="organ: prostate; Vector: pCMV-Sport6; site_1: NotI; site_2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NH_MGC library."
BASE COUNT		190 a 351 c 305 g 145 t
ORIGIN		
Query Match		36.0%; Score 682.6; DB 10; Length 991;

Best Local Similarity 95.6%; Pred. No. 1.2e-116;  
Matches 735; Conservative 0; Mismatches 29; Indels 5; Gaps 3;

OY	513	cgggggagctgtgtctcgtagagagaggacttgnatcccaatccaccccgacgcccagc	572
Db	1	CGGGGGAGTGGTGTCCAGAGAGGGAGACTGTGAACTCCAGTCCACCGGCAAGCCCGAC	60
OY	573	aggagctatccgctctcgacctccctcactcttgcacatgctgttacctgtagacagcgt	632
Db	61	AGGGCTTCCGGCTTCCTTACCTCTCTACTTGTGACATGTGTATCCCTGAGAGACAGACT	120
OY	633	gggcaagccaagggccccctgggggagcagctcgggagggagccactgcacagccttagcagt	652
Db	121	GGGGAGCCCAAGCCCCCTGTGGGCGCAGAGTCCGGGAGGAGCCACCTGAGAGACCTTAGAGAGT	180
OY	693	ggccggtatcttggagaagcccaagccccagcgggcacccttgagctgtgccccgggggtga	752
Db	181	GGCCGGTATTGTAGACAGCCCAAGCCCCCAGCGGGGCACTCTGGATTGGTATCCCGCGGGCTGA	240
OY	753	ccttgagagagcactgcgttgtagcaggtgtcagctccatggtgtgtgtgcgaagtctcaag	812
Db	241	CCTTGGAGGAGACTCGCTGGAGGACAGGTGCAGTCTCATGTGTGTGGGGAAGTGTCAAG	300
OY	813	acatcgagacgycctgcgaagtgtctcaacatcacgcgaagatcccaatgagcttggaccca	872
Db	301	ACATCGAGACGCGCCGCMAGCTGCTCAACATCACCGGACAGATCCCATGAGCTGAGCCCCA	360
OY	873	gcaatgtcacaagaatgtgtctctgttggagacagagcccaataaccgctgtcccccatggca	932
Db	361	GCAATGTCCACAAGTGGCTCTCTGTGGAGAGGACCAATATACCGGCTCCCGCCATGGGCA	420
OY	933	aggcctccagagagctgtgcggaggaagagctgtgcgcacatgtcgagagagcagttccgc	992
Db	421	AGGCTTCCAGAGACTGGCGGGGCAAGAGCTGTGCGCAATGTGGAGAGACAGTTCGGCC	480
OY	993	agcgctcgccccctgtgtgtggatgtgtctgcacgcgccaccttgacaatctggaatcagcgg	1052
Db	481	AGCGCTCGCCCCCTGGTGGGATGTGCTGCACGCCCAACTGTGACATCTGGAAGTCAACCG	540
OY	1053	cctgtgataaagagcgagcttcaccttggcgatccgattcactactgtgtctgaccaagttagg	1112
Db	541	CCTGTGATTAAGAGAGGGGCACTTCACTCTGGGGGATTCATCTATGTGCTCGACACAGTAGG	600
OY	1113	agagctgagcagacagcggaagtgtgacctatcatgtctccgggcaagcccatccactgtgac	1172
Db	601	AGAGCTGAGACCGACAGCCGAGTGAATCATATGCTCTCGGGCAACC--ATCCACTGTGGC	658
OY	1173	agttcctcaagagttgtctacttcaagcccccaagctatygccgtctcatagtgtgtca	1232
Db	659	AGTTCTCTCAAGAGTGTCTACTCAAG--CCCAAGCTATGAGCCGCTTATTAAG--TGCTC	715
OY	1233	acaaggaagaaggtcatcttcaaaattgagagctcaagcccaggtgtgccc	1281
Db	716	ACCAAGAAACGGCTCTTCCACATTTGAGACTCAAGCAAGCAAGGCGGCCCGG	764

RESULT	4
AI745526/c	
LOCUS	
DEFINITION	
AI745526	721 bp mRNA linear EST 17-DEC-1996
wc3hai10.x1 NCI_CGAP_Pr28	Homo sapiens cDNA clone IMAGE:2317050 3'
similar to SW:ETS4_DROME	P29775 DNA-BINDING PROTEIN D-ETS-4 ; , mRNA
sequence.	
AI745526	
AI745526.1	GI:5113814
EST.	
human.	
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
1 (bases 1 to 721)	
NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .	
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	
Tumor Gene Index	

JOURNAL  
COMMENT

unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: c9apbs-r@mail.nih.gov  
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www.bio.llnl.gov/bdrp/image/image.html  
 Insert length: 826 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 448.  
 Location/Qualifiers

FEATURES  
source

1. 721  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2317050"  
 /clone\_1lb="NCI\_CGAP\_Pr28"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)  
 with a modified polylinker; Plasmid DNA from the  
 normalized library NCI-CGAP Pr22 was prepared, and s  
 circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (clonoids  
 985608-986759, 1101192-1101959, and 1217928-1220615).  
 Subtraction by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 122 a 197 c 241 g 156 t 5 others  
 ORIGIN

Query Match 34.8%; Score 659.6; DB 9; Length 721;  
 Best Local Similarity 97.0%; Pred. No. 2e-112;  
 Matches 701; Conservative 0; Mismatches 19; Indels 3; Gaps 3;

OY 1172 agttccctcaaggaggtctactcaagcccaagcagctatgacgctcattagctgc 1231  
 DB 721 CAGTTCCTCAAGAGTTTGTCTACTC-AGCCCNACGATATGCCCTTCATAGAGTGCGTC 663  
 OY 1232 aacaaggaggaagcattcctcaaatgtgaagctcaagccagctgacgctgtgggc 1291  
 DB 662 ACMAAGAGAGAGGCGATC-TCMAAATTGAGGAGCTCAGCCAGGTGGCCCTGTGGGCG 604  
 OY 1292 atcccaagaagaccgtccgcagctgaactgagcaagctgacgctcaccgacagctc 1351  
 DB 603 ATCCGAGAGACCGTCCGCAATGACATAGCAAGCTGAGCCGCTCCATCCGCCAGT-A 545  
 OY 1352 tacaagaaggagcattcctcgaagcagacatctccagcgctcgtctaccagttcgtg 1411  
 DB 544 TACAGAAAGGCGATCATCGNAGCAACATCTCNCAGGCCCTCGTACCAAGTTGCTG 485  
 OY 1412 ccccccatgtggtccctcggccagggcctgaaccccgccagggcctctctctctgc 1471  
 DB 484 CACCCCATGTGAGTCCCTGCGCCAGGGCTTAACCGCCCTCAGGGGCTCTCTCTGCG 425  
 OY 1472 ctgagccctcagcagcgccctgagatggtggggaacacgggcaagtgtgtcctcgtcct 1531  
 DB 424 CTGCGCTGCTCAGCGACGGCCCTGAGATGGGGGAACGGGCACTGTCTGTGCTGT 365  
 OY 1532 gaccttcagagagcccaaggtcagggagggcaaccaactgcgccagggagatatggctc 1591  
 DB 364 GACCTTCCAGAGCCCAAGGTCAAGGAGGCGCAACCACTGCCCAAGGGGATATGGGTCG 305  
 OY 1592 tctgagccctcagcagcactgagggaggggtgtctcctcctcagggccagctcctccct 1651  
 DB 304 TCTGGGGCTTTCGGGACCTTGGGGAGGGGCTCTCTCTCTCAGAGCCCACTGTCTCCCT 245  
 OY 1652 gtagagcagagggagagcagggctgtctcccaacacctgtcctcgtacccagcattccag 1711

DB 244 GAGGAGACAGAGGAGAGAGGCGTCCGCCAACACCTGCTCTGACCCACATTTCCAG 185  
 OY 1712 agcagagcctaccagaaggagcagtgactgaaagaagccagagctcagcctctc 1771  
 DB 184 AGCAGACCTCAACAGAGGCGAGTCACTGACAAAGGCCAGCGAGTCCAGCCTCTCTC 125  
 OY 1772 tgcctcaccctcgtcctccattctgacacacactgcatgtgtcagagagactctg 1831  
 DB 124 TGCTCCATCCCTCCCTGCTCCATTTGACACACCTGGCATGTGCGAGGAGACATCTG 65  
 OY 1832 caccctcagttggcagcagagagtgccccgggaatgtgaataaataactacagaa 1891  
 DB 64 CACCCCTGATTTGGCAGCCAGAGTGCCTCCGGGAGATGATTAAGATCTTAACAA 5  
 OY 1892 ctg 1894  
 DB 4 CTG 2

RESULT 5  
 BF589055/c 640 bp mRNA linear EST 12-DEC-2000  
 LOCUS  
 DEFINITION  
 nacc2d04.x1 lupski\_sciatic\_nerve Homo sapiens cDNA clone  
 IMAGE:3394614 3' similar to TR:095238 095238 ETS TRANSCRIPTION  
 FACTOR PDEF.; mRNA sequence.  
 ACCESSION  
 BF589055.1 GI:11681379  
 KEYWORDS  
 EST.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 640)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: c9apbs-r@mail.nih.gov  
 Tissue Procurement: Dr. James Lupski  
 cDNA Library Preparation: Lupski Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 info@image.llnl.gov  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 416.  
 Location/Qualifiers

1. 640  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3394614"  
 /clone\_1lb="Lupski\_sciatic\_nerve"  
 /sex="male"  
 /tissue\_type="sciatic nerve"  
 /dev\_stage="adult", 70 yr  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORE6 (Life Technologies); Site-1:  
 NotI; Site-2: SalI; cDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCAGCGCGC-3' and  
 5'-CTGACTTCATGATCGCGAGGGGCGCCCT(15)-3'. Size selected >  
 1 kb for average insert length 1.87 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life  
 Technologies."

FEATURES  
source

BASE COUNT 103 a 182 c 218 g 136 t 1 others  
 ORIGIN









[illegible]

RESULT	10
BF446190/c	
LOCUS	7P33808 x1 625 bp mRNA linear EST 01-DEC-2000 similar to TR:O95238 O95238 ETS TRANSCRIPTION FACTOR PDEF,, mRNA
DEFINITION	
ACCESSION	sequence.
VERSION	BF446190
KEYWORDS	EST.
SOURCE	human,
ORGANISM	Homo sapiens

REFERENCE	1 (bases 1 to 625)
AUTHORS	NCI-CCAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LIND, send email to: [infoimage.lind.gov](mailto:infoimage.lind.gov)  
Seq. primer: -40UP from Gt100  
High quality sequence stop: 460.  
Location/Qualifiers  
I. .625

```

BASE COUNT      113 a      174 c      202 g      135 t
ORIGIN
"/clone="IMAGE:3647751"
"/clone_id="NCI.CGAP_Pr28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI.CGAP_Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneids
965608-986759, 1101192-1101959, and 1217928-1220615)."
Subtraction by Bento Soares and M. Fatima Bonaldo."
others

```

	Query Match	Score	DB	Length
	Best Local Similarity	97.1%	Pred. No. 4, 9e-97	625
	Matches	597	Conservative	0; Mismatches 17; Indels 1; Gaps 1;
Qy	1280	cgagctgaggagatccgcgaagaacggtccgcacatgaactcgcgcgaagctgacgcgctcc	1339	
Db	625	cggctgtgggggcatccgcgaanccggtccgcgcatgaactc-gccaaagctgacccctcc	567	
Qy	1340	accgcgcagctcttaagaagagatcaccgcgaagccagacatctcccgcgctctgc	1399	
Db	566	atcccccagatggttacaagaaggcattctccgaagccagacatctcccgccctctgc	507	

OY	1400	taaaagttcgtgaaccccaatcgaatgacctgcctgcgcacagagcctaaacccgcgcctcaagg	1459
Db	506	TACCAAGTTCGTGACCCCACTTTGATGCTTG6CCCAAG6GCTGTAAACCGCCTCAAG66	447
OY	1460	cctctcctcgtgcctgcctgcctcagcagagccctgagatgggaggaaaacagggcagtg	1519
Db	446	CCCTCCTCCGCTGCGCCCTGCCTCAGCCAG6GCTCTGAGATGGGGGAAAAAGGGCAATTTT	387
OY	1520	ctctgcgtcctctacaccttcagaagcccaaggtaagggaagggaacaaccaactccccaagg	1579
Db	386	CTTTGCTGCTTTACCTTCCAAAGGCCCAAGGTCAAGAGAGGGCAACCACTGCCCCAG66	327
OY	1580	ggatatagtgctcctctggggccttcgggaccaatggggaagggtgtcctcctcctcagccc	1639
Db	326	GGATATATGGGCTCTTTGGGGCTTCGGGACCTGGGGCAAG6GGTGTCTTCCTCCACAG6CC	267
OY	1640	agctgcctccccctggaggaagaagggagaagggctgtctccccaacacgtgcctctacac	1699
Db	266	AGCTGCTCCCTGTGAAGACAGAGGAGACAG6GGTGGCTCCCAACACCTGCTCTTATGCC	207
OY	1700	cagcatcttcagaagcagagcctacagaaggagggcagtgactcgacaagaagccacacgagc	1759
Db	206	CAGCATTTCCAGAGCAGAGCCTACAGAAAGGCACTGACTGCACAAAGGCCACAGGCAATC	147
OY	1760	caggcctctctcgtgcctccatcccccttgctcccaatctgcgaccaacactggcagtgtga	1819
Db	146	CAGGCTCTTTTGTGCTCATCCCTGCTCTCCCAATTTTGACACACACCTGGCAATGGTGA	87
OY	1820	gggagacatctgacccccctcagctgggcagcgcagagagtgcccccgagatglaaataa	1879
Db	86	GGGAGACATTTTGACCCCTGAGTTGGGACGACGAGAGTGCCTCCCGGGAATAGATATAA	27
OY	1880	gatactcagagaactg	1894
Db	26	GATCTAGAGAACTG	12

```

RESULT 11
BM006343 LOCUS 590 bp mRNA 11near EST 30-OCT-2001
DEFINITION 630615035F1 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5420913 5',
mRNA sequence.
ACCESSION BM006343
VERSION
KEYWORDS
SOURCE EST.
ORGANISM human.
Human.
Homosapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 590)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNT)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNT at:
http://image.llnl.gov
Plate: LLCM1876 row: e column: 10
High quality sequence stop: 578.
Location/Qualifiers
1..590
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5420913"
/clone_lib="NIH_MGC_110"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;"
FEATURES
Source

```

Site-2: EcoRI: cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(C). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 101 a 225 c 180 g 84 t  
ORIGIN

Query Match 30.3%; Score 573; DB 10; Length 590;  
Best Local Similarity 98.3%; Pred. No. 2.1e-96;

Matches 579; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 60 cagctctcccaagcctctgtcccaactccctgcgaagccctcaggttggtgacctgcagaggt 119  
DB 2 CAGTCTCTCAAGCCTGTGTGCAAGCTCTCCCTCAAGCCCTCAGAGTTGGCTTGCACAGGT 61  
QY 120 gccagcagcagcagccttggtggttaaggagactccctacagagcagcagccttgagac 179  
DB 62 GCCAGCAGGAGCCCTGTGGGTGGGGAGGGGAGCTCCCTAAGAGGAGCAGCCCTGTAGAC 121  
QY 180 gccagcagcagcagccttggtggttggtggtggtggtggtggtggtggtggtggtggt 239  
DB 122 CTGAGAGGGGACCCCTGTGAGGGTGGCCAGGCCCCAGTGGCCACCTGAGTGTCTCTC 181  
QY 240 tgcacacagcagcctgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 299  
DB 182 TGGCAGCAGCCCTGTGGGTGGGGAGGGGAGCTCCCTAAGAGGAGCAGCCCTGTAGAC 241  
QY 300 ccagtggtcctcagctgcacacactctccggcccttgagtggtggtggtggtggtggtggt 359  
DB 242 CCAAGTGGCCCTCAGCTGTGCAACCTCTTCCGCCCCCTAAGTGGCAGCTGAGCAGACA 301  
QY 360 gctccctgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 419  
DB 302 GCTTCCTGTGGGACAGGAGCTTACAGACAGCCGCGCAAGACAGAGAGGAGGATGG 361  
QY 420 gccagcagcagcagccttggtggtggtggtggtggtggtggtggtggtggtggtggtggt 479  
DB 362 GCAGCGCCAGCCCGGGGTCTGAGCAGGCTATCCGCCAGACACCTCTGAGACCCCGGACA 421  
QY 480 cgtgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 539  
DB 422 CGGTGTGCGGAGACGCTTGTGAGAGAGGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 481  
QY 540 acttgagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 599  
DB 482 ACTGAGAGTCCCAAGTCCAGCCGCGCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 541  
QY 600 acttgagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 648  
DB 542 ACTTTGACATGTGTACCTCTGAGGAGCAGCAGCTGTGGGAGCAAGGAGGAGGAGGAGGAG 590

RESULT 12  
LOCUS BG328411 899 bp mRNA linear EST 27-FEB-2001  
DEFINITION 60247779F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4546860 5',  
Mammalia sequence.  
ACCESSION BG328411  
VERSION BG328411.1 GI:13134849  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 899)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: g9abs-remail.nih.gov  
Tissue Procurement: ARCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LUCM1234 Row: b Column: 13  
High quality sequence stop: 602.

FEATURES  
source  
Location/Qualifiers  
1..899  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4546860"  
/clone\_id="NIH\_MGC\_15"  
/issue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: colon; Vector: pOTB7; Site:1: XhoI; Site-2:  
EcoRI: cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCAAGAG(C). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 163 a 346 c 277 g 113 t  
ORIGIN

Query Match 29.08; Score 549.8; DB 10; Length 899;  
Best Local Similarity 86.9%; Pred. No. 4.5e-92;  
Matches 734; Conservative 0; Mismatches 92; Indels 19; Gaps 11;

QY 5 gacttcctcccaagcagactctcgtcactctgctggttcacacagcagcagcagcagcagcagc 64  
DB 2 GACTTCCTCCAGCAGCATTTCTGCACTGTGCGGTGACACTGTGCCCAAGACAGACAGTGC 61  
QY 65 ctccaagcctgtctgcagctccctgcagcagcctcaggttggtggtggtggtggtggtggtggt 124  
DB 62 CTCCAAGCCTGTGCTGAGGCTCCCTGCAAGCCCTCAGGCTTGGGCTTGGCAGCGGTGCAAG 121  
QY 125 cagcagcagcctgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 184  
DB 122 CAGGCAAGCCTGTGCTGAGGCTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 181  
QY 185 ccaagcagcagcctgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 244  
DB 182 AGGGCCACACCTTTAGGCTGGCCAGGCGCCCAAGTGGCAGACAGTGTGCTTGTGCA 241  
QY 245 ccagcctgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 304  
DB 242 CCAGCCCTGTGCTGAGGCTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 301  
QY 305 ggcctcagctgagcagcagcctcttcgagcccttgaggttgagctgcagagagagagctcc 364  
DB 302 GGCTCAAGCTGCCACACACTTCTTCCGCGCCCTTGAATTTGGCACTGCAGACAGACAGTCC 361  
QY 365 ctggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 424  
DB 362 CTGGGACACAGGAGCTTACAGACAGAGTGG-CAGCCCAAAAGAGAGGCGCATGGGAGAG 420  
QY 425 gccagc 483  
DB 421 GCCAGCCCGGGCTGTGAGCAGCGATATCCCAAGCAGCAGCTCTGCTGCCCCACACAGG- 479  
QY 484 gtccgagc 542  
DB 480 GTCCGGGAGCAGGCTTGGAGAGAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 539  
QY 543 ggaatccagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 602  
DB 540 GGAATCCAGCTCAGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 594

BASE COUNT	ORIGIN
104 a	172 c 204 g 126 t

Query Match	29.0%	Score 549.6	DB 9	Length 606
Best Local Similarity	97.2%	Pred. No. 4.5e-92		
Matches 591	Conservative	0	Mismatches 14	Indels 3
				Gaps 3
QY 1288	ggagatccggaagaaacgctccgcacatgacgtgcagcaagctgtgacgctccatccgca	1347		
Db 606	GGGATGCCCAAGAA-GTGTCCGCCCTGAACTACGAACAGTGTGAGCGGCTCATTCGGCA	548		
QY 1348	gtcttaaaagaagagcatatccggaagcagacatctccagagcttgcttaacagt	1407		
Db 547	GTATTACAGAGAGGGCATCTATCCGGAGGCAACATTTTCCAGGCGCTCTTACCAAGTT	488		
QY 1408	cgtagaccccatctgtagtgccttggcccagggcttgaaccccgccctcaggggctctc	1467		
Db 487	CGTGCACCCCATCTGAGTGGCTGGCCACAGGGCTGTAACCCGGGCTCAGGGGCTCTGTC	428		
QY 1468	ctgagctcctctgcttaagcaagccctgtgagtgaggggaaacggtgctgtctgtcgt	1527		
Db 427	CTGGCTCTCCCTGCTTACGCCAGGCGCTCTGATGGGGGAAACGGGCGATGCTGCTGCTG	368		
QY 1528	ctctgaccttcacagagcccaaggtctcagggaggggcaacaactgcccacaggggata	1587		
Db 367	CTTGACACTTTCAGAGGCCCAAGTCAAGGAGGGGCAACCACTGCCACAGGGGATATGG	308		
QY 1588	gtccctctggggctctgggacatgggggaaggggtgtctctctcctcagccagctgtc	1647		
Db 307	GTCTCTGGGGG-CTTGGGAGCCTTGGGGAGGGGAGTCTCTCTCTGAGGCGAGCTGCTC	249		
QY 1648	ccctggaggaagaagggagagcaagggctgtccccaacacctgtcctctgacccagatt	1707		
Db 248	CCGTGGAGGAGACAGGGGAGACAGGGGCTGCTCCCAACACCTGCTCTGACCCAGCATTT	189		
QY 1708	ccagagcagagagccttaagaagggcagtgactctgcacaagccacagggcagtcacgctc	1767		
Db 188	CCAAAGAGAGAGCCTTACAGAAAGGGCAGGTGCTGACAAAGAGCCACAGCACTTCAGGCTTC	129		
QY 1768	tctctgtcatcattccctcctcctccattctgtgacacacacctgtgactgttcagggagca	1827		
Db 128	TCTGTGTCATCTCCCTGCTGCTCCATTTCTGTGACACACCTGGCATGTCTCAGGAGATA	69		
QY 1828	tctgacccctcagttggagcagccagagtgccccg-ggaatgataataagaatacta	1886		
Db 68	TCTCTACCCCTGAGTTGTGCAGCGCATGAGTGCCCGCCGGTAATGATAAATAAGATACTA	9		
QY 1887	gagagactg 1894			
Db 8	GAGAACTG 1			

RESULT 14

AW130657/c 652 bp mRNA linear EST 27-OCT-1999

LOCUS x664607.x1 NCL CGAP UT3 Homo sapiens cDNA clone IMAGE:2612652.3, similar to TR:095238.095238 ERS TRANSCRIPTION FACTOR PDEF. ;, mRNA sequence.

ACCESSION AW130657

VERSION AW130657

KEYWORDS EST.

SOURCE AW130657.1 GI:6132353

ORGANISM human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R. Emmerit-Buck, M.D., Ph.D.

CDNA library preparation: Life Technologies, Inc.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.lnl.gov/bbrp/image/image.html](http://www.bio.lnl.gov/bbrp/image/image.html)  
 Seq primer: -400p from Gibco  
 High quality sequence stop: 423.

## FEATURES

source

Location/Qualifiers

1..652  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2612652"  
 /clone.lib="NCI CGAP. ut3"  
 /tissue.type="poorly-differentiated endometrial  
 adenocarcinoma; 2 pooled tumors"  
 /lab\_host="DH10B"  
 /note="Organ: uterus; Vector: PCMV-SPORT6; Site\_1: Sall;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.45 kb. life Technologies catalog #:  
 11541-018"

BASE COUNT 116 a 182 c 210 g 142 t 2 others

## ORIGIN

Query Match 27.8%; Score 525.8; DB 9; Length 652;  
 Best Local Similarity 94.9%; Pred. No. 1.2e-87;  
 Matches 564; Conservative 0; Mismatches 28; Indels 2; Gaps 2;

QY 1302 accgtcccccatagaactcgaacagctcgaagcgtcattccgcagagcttaacaagaagg 1361  
 |||||||  
 DB 600 ACCGTCCTCCGCAATGATACGACAACTTACCCG--TCAATCCGCAATATACAAAGAGG 542  
 QY 1362 gcatcctccggaagccagacatctccagcgtctcgtacacagtctgaccccatct 1421  
 |||||||  
 DB 541 GCATCATCGGAAGCCAGACATCTCCAGCGCTCTGCTACAGATTCGTGACACCCCATCT 482  
 QY 1422 gaatgctgcccagagcgttgaaacccgctcagggcgtctctcgtcgtcgtcc 1481  
 |||||||  
 DB 481 GAATGCTGCCCCAGGCTTGAAACCCGCCCTCAAGGGGCTCTCTGCTGCCCCCTCC 422  
 QY 1482 tcagcagagccctgaagatgagggaacagcagtgctcgtcgtcgtcgtcgtcgtc 1541  
 |||||||  
 DB 421 TCAGCAGGCGCTGAGATGGGGGGAANAGCGGCGAGTCTGTGCTGCTGACCTTCCAG 362  
 QY 1542 agcccaagctcagagagggaggaacacactgccccaggggagatggttcctctgggct 1601  
 |||||||  
 DB 361 AGCCCAAGTCAAGGAGGAGGCAACCAATGCCCAAGGGGAGATGGTCTCTGGGGCT 302  
 QY 1602 tcggagcaatggagcaggggtctctcctcagagccagctgctcccttgagagaca 1661  
 |||||||  
 DB 301 TCGGAGACCTGGGGGCAAGGGGTCTCTCTCTCAGGCCCAAGCTGCTCCCTGAGAGACA 242  
 QY 1662 gggagagagggctgctcctcccaacactgctcgtcgtcgtcgtcgtcgtcgtcgtc 1720  
 |||||||  
 DB 241 GGGAGACAGGGCTGCTCCCAACACTGCTCTGACCCAGCATTTTTCAGAGAGAGACC 182  
 QY 1721 tagaagagggcagtgactcgaacaaagggcagagcagtgctcgtcgtcgtcgtcgtc 1780  
 |||||||  
 DB 181 TAGAGAGGGCAATGACTGACAAAGGCCACAGGAGCTCAAGGCTCTCTCTGCTCCATC 122  
 QY 1781 cccctgctcccaatctgacccaacactgagctggtgagagggagagacatctgacccctca 1840  
 |||||||  
 DB 121 CCCCTGCTCCCATTTTGGACCACTGCTGAGTGTGAGGAGAGCAATCTGACACCTCTGA 62  
 QY 1841 gttaggagcagcagagtgccccggagagatgagataaagatactcagagagaatt 1894  
 |||||||  
 DB 61 GTTGGGAGCCAGAGACTGCCCGGGAATGATTAAGATCTAGAGAACTG 8

RESULT 15  
 AA662204  
 LOCUS AA662204 514 bp mRNA linear EST 03-DEC-1997  
 DEFINITION ns66c07.r1 NCI\_CGAP\_Pr22 Homo sapiens cDNA clone IMAGE:1188588 5'

similar to SW:ETS4\_DROME P29775 DNA-BINDING PROTEIN D-ETS-4; , mRNA  
 sequence.  
 AA662204  
 VERSION AA662204.1 GI:2616295  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.lnl.gov/bbrp/image/image.html](http://www.bio.lnl.gov/bbrp/image/image.html)  
 Insert Length: 734 Std Error: 0.00  
 Seq primer: -28m13 rev1 ET from Amersham.

## FEATURES

source

Location/Qualifiers

1..514  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1188588"  
 /clone.lib="NCI CGAP\_Pr22"  
 /sex="male"  
 /tissue.type="normal prostate"  
 /lab\_host="DH10B"

/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)  
 with a modified polylinker; 1st strand cDNA was prepared  
 from normal prostate bulk tissue, and was then primed with  
 a Not I - oligo(dT) primer. Double-stranded cDNA was  
 ligated to Eco RI adaptors (Pharmacia), digested with Not  
 I and cloned into the Not I and Eco RI sites of the  
 modified pT73 vector. Library is normalized, and was  
 constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 108 a 173 c 151 g 82 t

## ORIGIN

Query Match 26.6%; Score 504.4; DB 9; Length 514;  
 Best Local Similarity 98.8%; Pred. No. 1e-83;  
 Matches 508; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1240 gaaggcatcttcaaatgaggaactcagccagtgcccgctgtgaggacatcgcaa 1299  
 |||||||  
 DB 1 GGAGGGCATCTTCAAAATGAGACACACCCAGAGGCGCCGCTGTGGGCATCCGCA 60  
 QY 1300 gaacgtcccgcatgaactgagcagctgagcgtcctcagtcgagcttaacaaga 1359  
 |||||||  
 DB 61 GAACGTCGCCGATGATACTACGACAAAGCTGAGCCGCTCATCCGCAATTAAGAA 120  
 QY 1360 gggcatcatcctggaagccagacatctccagcgcctgctcctcagcttgtagcccat 1419  
 |||||||  
 DB 121 GGGCATCATCCGGAGACCCAGACATCTCCAGCGGCTGCTACGATTCGTCACCCCAT 180  
 QY 1420 ctgagtgcctgagccagagggcctgaacccgcctcagagggagctcctcgtcgtcgtc 1479  
 |||||||  
 DB 181 CTGAGTCCCTGGACCAAGGGCATGAACCCGCCCTCAAGSGGCTCTCTCTGCTCCCTG 240  
 QY 1480 cctcagcagggccttgagatgagggaacacgggcaagtgtgtctcgtcgtcgtcgtc 1539  
 |||||||  
 DB 241 CCTACGCAAGGCTGAGATGGGGGAAAGGGCACTGTCTCTGCTGACACTTCC 300  
 QY 1540 aagggccaagtgtagggaggggacacaaatgccccagggagatggtctcctcggggc 1599  
 |||||||



Db 301 AGAGCCCAAGGTCAAGGAGGGGCAACCAACTGCCCCAGGGGATATGGGTCTCTGGGGC 360  
QY 1600 ctctgggacacatgggcaagggtgtctctcctcctcagggccagctgtcccttgaggaca 1659  
Db 361 CTTCGGGACCATGGGCGAGGGGTGTCTCTCTCTCAGGCCCCAGCTGTCTCCCTGGAGACA 420  
QY 1660 gaggagacaggggtgtctcccaacacctgtccttgacccagcatcttcagagcagagc 1719  
Db 421 GAGGGAAGACAGGGCTGCTCCCAACACCTGTCTGTGACCCCAACATTTCCAGAGCAGAGC 480  
QY 1720 ctacagaaggcagtgactcgacaaaggccacag 1753  
Db 481 CTACAGAAGGGCACTGACTCGACAAAGGCCACAG 514

Search completed: August 6, 2002, 18:52:17  
Job time: 2286 sec



Db 301 cagtgagctcagctgcacacactcttcctccgctgaagtgtgacgtgcaagacag 360  
Qy 361 ctcccttgagcagcagagctaaacagacagcagccagccaaacagcagcgtgag 420  
Db 361 ctcccttgagcagcagagctaaacagacagcagccagccaaacagcagcgtgag 420  
Qy 421 cagcgacagcccggtctgagcagcgtatcccccagccactctctgcccccgacac 480  
Db 421 cagcgacagcccggtctgagcagcgtatcccccagccactctctgcccccgacac 480  
Qy 481 ggtgtctgagcagcagcttgagagagcagcagcagcagcagcagcagcagcag 540  
Db 481 ggtgtctgagcagcagcttgagagagcagcagcagcagcagcagcagcagcag 540  
Qy 541 ctgagagctccagctccacccagccagccagcagcagcagcagcagcagcagcag 600  
Db 541 ctgagagctccagctccacccagccagccagcagcagcagcagcagcagcagcag 600  
Qy 601 ctctgacacatgctgacccctgagagcagcagcagcagcagcagcagcagcag 660  
Db 601 ctctgacacatgctgacccctgagagcagcagcagcagcagcagcagcagcag 660  
Qy 661 tcggagagagccacctcagcagcagcagcagcagcagcagcagcagcagcagcag 720  
Db 661 tcggagagagccacctcagcagcagcagcagcagcagcagcagcagcagcagcag 720  
Qy 721 gggcagacccctgagctgagcagcagcagcagcagcagcagcagcagcagcag 780  
Db 721 gggcagacccctgagctgagcagcagcagcagcagcagcagcagcagcagcag 780  
Qy 781 gcaatcctcagctgagcagcagcagcagcagcagcagcagcagcagcagcagc 840  
Db 781 gcaatcctcagctgagcagcagcagcagcagcagcagcagcagcagcagcagc 840  
Qy 841 calcacccagcagatccacatgagcagcagcagcagcagcagcagcagcagcagc 900  
Db 841 calcacccagcagatccacatgagcagcagcagcagcagcagcagcagcagcagc 900  
Qy 901 agagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 960  
Db 901 agagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 960  
Qy 961 gctgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1020  
Db 961 gctgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1020  
Qy 1021 gcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080  
Db 1021 gcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080  
Qy 1081 ggcagatcactactgctgctcagcagcagcagcagcagcagcagcagcagcag 1140  
Db 1081 ggcagatcactactgctgctcagcagcagcagcagcagcagcagcagcagcag 1140  
Qy 1141 atcatgtctccgggagcagcagcagcagcagcagcagcagcagcagcagcagc 1200  
Db 1141 atcatgtctccgggagcagcagcagcagcagcagcagcagcagcagcagcagc 1200  
Qy 1201 ccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260  
Db 1201 ccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260  
Qy 1261 ggaactcagccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
Db 1261 ggaactcagccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
Qy 1321 cgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380  
Db 1321 cgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380  
Qy 1381 catctccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440  
Db 1381 catctccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440

Qy 1441 tgaacccgcctcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1500  
Db 1441 tgaacccgcctcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1500  
Qy 1501 ggggaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1560  
Db 1501 ggggaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1560  
Qy 1561 ggaacacactgcccagcagcagcagcagcagcagcagcagcagcagcagcag 1620  
Db 1561 ggaacacactgcccagcagcagcagcagcagcagcagcagcagcagcagcag 1620  
Qy 1621 gtcctctctcctcagcagcagcagcagcagcagcagcagcagcagcagcag 1680  
Db 1621 gtcctctctcctcagcagcagcagcagcagcagcagcagcagcagcagcag 1680  
Qy 1681 caacacctgctctcagcagcagcagcagcagcagcagcagcagcagcagcag 1740  
Db 1681 caacacctgctctcagcagcagcagcagcagcagcagcagcagcagcagcag 1740  
Qy 1741 acaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1800  
Db 1741 acaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1800  
Qy 1801 ccacacctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1860  
Db 1801 ccacacctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1860  
Qy 1861 cccgggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1894  
Db 1861 cccgggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1894

RESULT 2  
US-10-172-118-1406  
Sequence 1406, Application US/10172118  
GENERAL INFORMATION:  
APPLICANT: Dai, Hongyue  
APPLICANT: He, Yudong  
APPLICANT: Linsley, Peter  
APPLICANT: Mao, Mao  
APPLICANT: Roberts, Chris  
APPLICANT: Van 't Veer, Laura  
APPLICANT: Van de Vijver, Marc  
APPLICANT: Bernards, Rene  
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
FILE REFERENCE: 9301-175-999  
CURRENT APPLICATION NUMBER: US/10/172,118  
CURRENT FILING DATE: 2002-06-14  
PRIOR APPLICATION NUMBER: 60/380,770  
PRIOR FILING DATE: 2002-05-14  
NUMBER OF SEQ ID NOS: 2699  
SEQ ID NO 1406  
LENGTH: 1894  
TYPE: DNA  
ORGANISM: Homo sapiens  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: NM\_012391  
DATABASE ENTRY DATE: 2001-06-18  
US-10-172-118-1406

Query Match 98.7%; Score 1870; DB 7; Length 1894;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1879; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 gctgtactctctccagcagcagcagcagcagcagcagcagcagcagcagcagcag 60  
Db 1 gctgtactctctccagcagcagcagcagcagcagcagcagcagcagcagcagcag 60  
Qy 61 agtctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 120  
Db 61 agtctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 120











```

: PRIOR FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: 60/325,102
: PRIOR FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 60/323,580
: PRIOR FILING DATE: 2001-09-19
: NUMBER OF SEQ ID NOS: 363
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 74
: LENGTH: 1907
: TYPE: DNA
: ORGANISM: Homo sapiens
: PCT-US02-07826-74

```

```

Query Match          4.0%; Score 75; DB 1; Length 1907;
Best Local Similarity 59.0%; Pred. No. 2.8e-05;
Matches 148; Conservative 0; Mismatches 100; Indels 3; Gaps 1;

```

```

QY 1159 catccacctgtgagcttctcaaggagttgctactcaagcccaagctatgcccgtt 1218
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 911 caccacccctgtgagcttctcaaggagttgctactcaagcccaagctatgcccgtt 970

QY 1219 cataggtgtgcacaagaagagagagcattcaaatgagagctcagcccaggtgac 1278
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 971 catgaagtgagagagagagagagagagagagagagagagagagagagagagagag 1030

QY 1279 ccgagctgtgagagagagagagagagagagagagagagagagagagagagagag 1338
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1031 ccaactatgtgggccaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1090

QY 1339 catccgcagcttcaagaagaagagagagagagagagagagagagagagagagagag 1398
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1091 catgag--gtactactacaagaagagagagagagagagagagagagagagagagagag 1147

QY 1399 ctaccagctcg 1409
   ||||| |||||
Db 1148 ctacaagtttg 1158

```

RESULT 9  
US-10-097-340-74

: Sequence 74, Application US/10097340

: GENERAL INFORMATION:

: APPLICANT: John MONAHAN

: APPLICANT: Manjula GANNAPURU

: APPLICANT: Sebastian HOERSCH

: APPLICANT: Shubhangi KAMATKAR

: APPLICANT: Steve G. KOVATS

: APPLICANT: Rachel E. MEYERS

: APPLICANT: Michael MORRISSEY

: APPLICANT: Peter OLANDT

: APPLICANT: Aml SEN

: APPLICANT: Peter VEIBY

: APPLICANT: Gordon B. MILLS

: APPLICANT: Robert C. BAST, Jr.

: APPLICANT: Karen LU

: APPLICANT: Rosemarie SCHMANT

: APPLICANT: Xumei ZHAO

: APPLICANT: Karen GIATT

: TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,

: TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer

: FILE REFERENCE: MRI-030

: CURRENT APPLICATION NUMBER: US/10/097,340

: CURRENT FILING DATE: 2002-03-14

: PRIOR APPLICATION NUMBER: 60/276,025

: PRIOR FILING DATE: 2001-03-14

: PRIOR APPLICATION NUMBER: 60/325,149

: PRIOR FILING DATE: 2001-09-26

: PRIOR APPLICATION NUMBER: 60/276,026

: PRIOR FILING DATE: 2001-03-14

: PRIOR APPLICATION NUMBER: 60/324,967

: PRIOR FILING DATE: 2001/09/26

: PRIOR APPLICATION NUMBER: 60/311,732

```

: PRIOR FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: 60/325,102
: PRIOR FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 60/323,580
: PRIOR FILING DATE: 2001-09-19
: NUMBER OF SEQ ID NOS: 363
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 74
: LENGTH: 1907
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-097-340-74

```

```

Query Match          4.0%; Score 75; DB 7; Length 1907;
Best Local Similarity 59.0%; Pred. No. 2.8e-05;
Matches 148; Conservative 0; Mismatches 100; Indels 3; Gaps 1;

```

```

QY 1159 catccacctgtgagcttctcaaggagttgctactcaagcccaagctatgcccgtt 1218
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 911 caccacccctgtgagcttctcaaggagttgctactcaagcccaagctatgcccgtt 970

QY 1219 cataggtgtgcacaagaagagagagagagagagagagagagagagagagagagag 1278
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 971 catgaagtgagagagagagagagagagagagagagagagagagagagagagagag 1030

QY 1279 ccgagctgtgagagagagagagagagagagagagagagagagagagagagagag 1338
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1031 ccaactatgtgggccaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1090

QY 1339 catccgcagcttcaagaagaagagagagagagagagagagagagagagagagagag 1398
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1091 catgag--gtactactacaagaagagagagagagagagagagagagagagagagagag 1147

QY 1399 ctaccagctcg 1409
   ||||| |||||
Db 1148 ctacaagtttg 1158

```

RESULT 10  
US-09-919-002-13047

: Sequence 13047, Application US/09919002

: GENERAL INFORMATION:

: APPLICANT: DeshKowitz, Dena

: APPLICANT: Liu, Jin

: TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA

: FILE REFERENCE: 20411-752CON1

: CURRENT APPLICATION NUMBER: US/09/919,002

: CURRENT FILING DATE: 2001-07-30

: PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922

: PRIOR FILING DATE: FILING DATE: 1999-07-22

: PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/024,341

: PRIOR FILING DATE: FILING DATE: 1998-02-13

: NUMBER OF SEQ ID NOS: 13203

: SOFTWARE: FastSeq for Windows Version 3.0

: SEQ ID NO 13047

: LENGTH: 2457

: TYPE: DNA

: ORGANISM: Homo sapiens

: US-09-919-002-13047

```

Query Match          4.0%; Score 75; DB 5; Length 2457;
Best Local Similarity 59.0%; Pred. No. 2.8e-05;
Matches 148; Conservative 0; Mismatches 100; Indels 3; Gaps 1;

```

```

QY 1159 catccacctgtgagcttctcaaggagttgctactcaagcccaagctatgcccgtt 1218
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 940 caccacccctgtgagagagagagagagagagagagagagagagagagagagagagagag 999

QY 1219 cataggtgtgcacaagaagagagagagagagagagagagagagagagagagagag 1278
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

[illegible][illegible]

```
TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2064 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1386
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-167-485-1
```

Query Match 3.9%; Score 74.8; DB 6; Length 2064;

Best Local Similarity 74.6%; Pred. No. 3e-05; Mismatches 32; Indels 0; Gaps 0;

```
QY 1250 ttcaaatgagactagccagcggtgcccgtctgtgggcatccgcaagaacgctccc 1309
DB 1039 ttcaagctcattgagcttgcagagctgcgacgctctggggcatccagaacgctcca 1098
QY 1310 gccatgacgagcagcaagcgtgacgctcattccgcaagcttaagaaggcctatcc 1369
DB 1099 gccatgatttaccagacagctgagcgcctcctccgattattatgagaaaggcctatcg 1158
QY 1370 cgggaag 1375
DB 1159 cagagag 1164
```

```
RESULT 14
US-10-099-926-1788
; Sequence 1788, Application US/10099926
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547c2
; CURRENT APPLICATION NUMBER: US/10/099,926
; CURRENT FILING DATE: 2002-03-17
; NUMBER OF SEQ ID NOS: 1982
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1788
; LENGTH: 2333
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-099-926-1788
```

Query Match 3.9%; Score 74.8; DB 7; Length 2333;  
Best Local Similarity 74.6%; Pred. No. 3e-05; Mismatches 32; Indels 0; Gaps 0;

```
QY 1250 ttcaaatgagactagccagcggtgcccgtctgtgggcatccgcaagaacgctccc 1309
DB 1308 ttcaagctcattgagcttgcagagctgcgacgctctggggcatccagaagaacgctcca 1367
QY 1310 gccatgacgagcagcaagcgtgacgctcattccgcaagcttacaagaaggcctatcc 1369
DB 1368 gccatgatttaccagacagctgagcgcctcctccgattattatgagaaaggcctatcg 1427
QY 1370 cgggaag 1375
DB 1428 cagagag 1433
```

RESULT 15

```
PCT-US02-12378-282
; Sequence 282, Application PCT/US0212378
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margareta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.4703PC
; CURRENT APPLICATION NUMBER: PCT/US02/12378
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-12378-282
```

Query Match 3.9%; Score 74; DB 1; Length 502;  
Best Local Similarity 59.5%; Pred. No. 4.3e-05; Mismatches 85; Indels 0; Gaps 0;

```
QY 1159 catccactgtgagcgttccctcaagaagtgctactcaagcccaacagctagccgctt 1218
DB 269 caccacccgtgagagcttccatccggtgacatctccaccggagctcaacgagggcct 328
QY 1219 catgagtgctcaacaagaggaagggcatcttcaaatgtaggagctcagccagctggc 1278
DB 329 catgaagtgagggaatcgcagatgaagggctctcaagtgctcgtcgcagagctggc 388
QY 1279 ccggtctgtgggcatccgcaagaacgctccgcgcatgactcgacaagctgaagcgtc 1338
DB 389 ccaactatgggccaacaaagaacaaagaacagacatgactcagagaagctgagcgggc 448
QY 1339 catccgcaagcttacaagaaggcctat 1368
DB 449 catgagttactactacaacgggagatcct 478
```

Search completed: August 6, 2002, 20:55:39  
Job time: 9463 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 18:16:51 ; Search time 87.61 Seconds

(Without alignments)  
5310.237 Million cell updates/sec

Title: US-09-841-963A-1

Perfect score: 1894  
Sequence: 1 gttctgacttcctccagcagc.....ataaagatactagagaactg 1894

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, NA: \*  
1: /cgn2\_6/ptodata/2/lna/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/lna/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/lna/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/lna/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/lna/PCrUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/lna/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1861.4	98.3	1905	4 US-09-055-113-2	Sequence 2, Appli
2	263	13.9	278	4 US-09-439-313-398	Sequence 398, App
3	83.2	4.4	2266	2 US-09-213-767-1	Sequence 1, Appli
4	75.2	4.0	2667	2 US-08-469-412A-1	Sequence 1, Appli
5	75.2	4.0	2667	4 US-09-021-715-1	Sequence 1, Appli
6	74.8	3.9	2064	3 US-08-875-944B-1	Sequence 1, Appli
7	74.8	3.9	2064	4 US-09-116-049-3	Sequence 3, Appli
8	73.4	3.9	1920	1 US-08-746-789A-1	Sequence 1, Appli
9	71	3.7	2938	2 US-08-343-443B-3	Sequence 1, Appli
10	70.2	3.7	2410	2 US-08-760-835B-1	Sequence 3, Appli
11	70.2	3.7	2410	4 US-09-303-268-1	Sequence 1, Appli
12	70.2	3.7	2410	4 US-09-116-049-1	Sequence 1, Appli
13	64.8	3.4	1752	3 US-09-360-779-1	Sequence 1, Appli
14	63	3.3	2268	3 US-09-344-579-1	Sequence 1, Appli
15	58.4	3.1	2544	2 US-08-469-412A-6	Sequence 6, Appli
16	58.4	3.1	2544	4 US-09-021-715-6	Sequence 6, Appli
17	57.4	3.0	852	4 US-09-020-956-44	Sequence 44, Appli
18	57.4	3.0	852	4 US-09-030-607-44	Sequence 44, Appli
19	57.4	3.0	852	4 US-09-439-313-44	Sequence 44, Appli
20	57.4	3.0	5427	3 US-09-009-913-2	Sequence 2, Appli
21	57.4	3.0	5510	3 US-09-009-913-3	Sequence 3, Appli
22	57.4	3.0	5667	3 US-09-009-913-4	Sequence 4, Appli
23	57.2	3.0	2975	1 US-08-368-281-1	Sequence 1, Appli
24	57.2	3.0	3240	1 US-08-368-281-3	Sequence 3, Appli
25	56.4	3.0	1604	1 US-08-306-691B-43	Sequence 43, Appli
26	56.4	3.0	1604	5 PCT-US93-06251-9	Sequence 9, Appli
27	54.8	2.9	848	3 US-09-009-913-338	Sequence 338, App

#### ALIGNMENTS

28	50.2	2.7	1447	4 US-08-878-177-1	Sequence 1, Appli
29	50.2	2.7	1528	4 US-08-878-177-3	Sequence 3, Appli
30	46.2	2.4	1035	1 US-07-601-094-30	Sequence 30, Appli
31	46.2	2.4	1035	1 US-08-012-735-30	Sequence 30, Appli
32	46.2	2.4	1281	4 US-08-998-416-3	Sequence 3, Appli
33	46.2	2.4	1910	4 US-09-593-711A-3	Sequence 3, Appli
34	46.2	2.4	1914	1 US-07-601-094-1	Sequence 1, Appli
35	46.2	2.4	1914	1 US-08-012-735-1	Sequence 1, Appli
36	45.8	2.4	1364	1 US-08-306-691B-50	Sequence 50, Appli
37	45.8	2.4	1364	5 PCT-US93-06251-65	Sequence 65, Appli
38	45.2	2.4	152331	3 US-09-128-155-16	Sequence 16, Appli
39	44.6	2.4	985	4 US-09-056-556-182	Sequence 182, App
40	44	2.3	23673	4 US-09-773-816-1	Sequence 1, Appli
41	43.6	2.3	15872	4 US-09-105-537-1	Sequence 1, Appli
42	43.6	2.3	43280	2 US-08-804-227C-1	Sequence 1, Appli
43	43	2.3	19182	2 US-08-850-880-11	Sequence 11, Appli
44	43	2.3	19182	2 US-08-944-916-11	Sequence 11, Appli
45	42.6	2.2	1227	4 US-09-385-028-23	Sequence 23, Appli

RESULT 1  
US-09-055-113-2  
; Sequence 2, Application US/09055113  
; Patent No. 6265565  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: PROSTATE-ASSOCIATED ETS PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
ADDRESS: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: PASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/055,113  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0501 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1905 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PROSTUT12  
CLONE: 1813005  
US-09-055-113-2  
Query Match 98.3%; Score 1861.4; DB 4; Length 1905;

Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1871; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 gcttgacttctccagacatctcgcactcgtccgtgtccacatgtcccaaaccc 60  
Db 15 gtttgacttctccacacacattctgacactgctggtgtccacactgccccacacacc 74  
QY 61 agtctccaaagcctgtctgcagctccctcgaagccctcaggttggccttgcacggtg 120  
Db 75 AGTCTCACAAGCCTGGCTGGCAGCTCCCTCGAAGCCCTCAGATTGGGCTTGCAAGG 134  
QY 121 ccaagcagcagcccttggtgtggtgtaaggagctccctcagcagcagccctgagcc 180  
Db 135 CCAGCAGGCGCCCTGGGCTGGGGGTAGGGGACTCCCTCAGCAGCAGCCCTTGAGAC 194  
QY 181 gcaaccaagccacccttgaagggtgtgcagagcccccagttgccaacttgatgtgctct 240  
Db 195 TCAGAGAGGCGCACCCCTTTGAGAGGtggccAGGCCCCCAGTGCCCAACTGAGTGCCTCT 254  
QY 241 gcaaccaagcctgtctgcccctgtctcgtctggtcccccaagatgcttggctgagacagc 300  
Db 255 GCCACGAGCCCTGCTGGGCCCTGTTCGGCTGGCCCCCAGATGCTTGCTGAGACAGC 314  
QY 301 caatggccttaagctgtgcacaaacctcttccggcccttgagatgtgcactgacagcagag 360  
Db 315 CAGTGGCTCTAGCTGCCACACACTCTCCGGCCCTTGAAAGTGGACATGACACAGACAG 374  
QY 361 ctcccttggagcagcagcagctaaacagacagcagccagcccaagcagcagctgag 420  
Db 375 CTCCCTGGGACACAGCAGCTAACACACAGCCGCGACGCCCAAGACAGCAGCGCATGGG 434  
QY 421 cagcgcagcccggtctgtgagcagctatcccccagcaacctctgtgtccccccagac 480  
Db 435 CACGCGCAGCCCGGGCTGTGAGCGCTATGCCCAAGCCTCTCTGCCCCGACACAC 494  
QY 481 ggtgtcggagcagcagcttggagagagcagcagcagcagcagcagcagcagcagcagc 540  
Db 495 GGTGTCCGGGACAGGCTTGGAGAGGGGGAGGGGGGAGAGTGGCTGTGAGAGAGAGGGGA 554  
QY 541 ctggagctcccaagctcaaccgcccaagcccgagcagcagcagcagcagcagcagcagc 600  
Db 555 CTGGAGTCCCAAGTTCACACGCCGCCAGCGGCCGCTGTCCGCTTACTCTCTCCTA 614  
QY 601 ctttgaacatgtcttaacctgaagcagcagcagcagcagcagcagcagcagcagcagc 660  
Db 615 CTTTGAATGTCTTACCTGTAGGACAGCAGCTGGGCAAGGCCCTCTGGGGCCAGCAG 674  
QY 661 tggggagggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
Db 675 TCGGGAGAGCCACCTGAGAGCCCTGACAGTGCCTGGCTCATTTGACAGCCCAAGCCCCAGC 734  
QY 721 gggcagcccttggacttgggtgcccggcggtgagcagcagcagcagcagcagcagcagc 780  
Db 735 GGGCAGCCTGAGCTTGGTGGCCCGGGCGGCTGACCTTGGAGAGACACTCCCTGGAGCAGGT 794  
QY 781 gcaagtcacatgtgtgtggtgagagtgctcaagaagacacagcagcagcagcagcagc 840  
Db 795 GCAAGTCCATGTGTGGTGGGCAAGTGTCTAAGGACATCGAGCGGCTGCAAGTGTCTCA 854  
QY 841 catcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900  
Db 855 CATCACCGCGAGATCCATGGACTGGAGGCCCAAGCAGCATGTGCAAGATGGCTCTCTGGAGC 914  
QY 901 agagcacaacatccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960  
Db 915 AGAGCAGCAGCATACCGGCTGGCCCCCATGGGCAAGGCGCTTCAGAGACTGGCGGCAAGGA 974  
QY 961 gctgtgtgcacatgtcggagagcagcagcagcagcagcagcagcagcagcagcagcagc 1020  
Db 975 GCTGTGTGGCATGTCTGGAGAGAGTTCGGCCAGCCTCTGCCCTTGGGTGGGATGTGCT 1034  
QY 1021 gcaagccacaccttggacatcaggagcagcagcagcagcagcagcagcagcagcagcagc 1080

Db 1035 GCACGCCACCTGAGCATGTGGAAGTCAAGCGGCTGTGATGAAGAGCGGACTTCACCTGG 1094  
QY 1081 ggcagatcaactactgtgtctgcagcagcagcagcagcagcagcagcagcagcagcagc 1140  
Db 1095 GCGATTACTACTGTGCTGTGACCAAGTGAAGAGAGTGGACCTGGACCGAGGTGGACTC 1154  
QY 1141 atcatgtctcggggagcccatccacactgtgtgcaatgtctcccaaggaggtgtcactcaagcc 1200  
Db 1155 ATCATGTCTCGGGAGCGCCATTCACACTGTGGCAGTTCCTCAAGAGGTGTACTCAAGGCC 1214  
QY 1201 ccaacagctatgtgcgcttcatatggtgtgtcacaagaagagagcagcagcagcagcagc 1260  
Db 1215 CCACAGCTATGGCCGCTTCAATTAGGTGCTCAACAGAGAGAGGCGCATCTTCAAAATTGA 1274  
QY 1261 ggaactcaagccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
Db 1275 GGACTCAGCCCAAGGTGGCGCCGCTGTGGGGCATCCGCAAGAACGCTCCCGCATGAACTA 1334  
QY 1321 cgaacagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380  
Db 1335 CGACAAAGCTGAGCCGCTCATCCGCAAGTATACAAAGAGGCGATTCGGAAGCCAGA 1394  
QY 1381 catctcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440  
Db 1395 CATCTCCAGCGCCCTGTCTACCAAGTGTGTCACCCCAATTCGATGCTTGCCCAAGGCC 1454  
QY 1441 tgaacccgcccctcaggggagcagcagcagcagcagcagcagcagcagcagcagcagc 1500  
Db 1455 TGAACCCGCCCTCAGGGGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1514  
QY 1501 ggggaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1560  
Db 1515 GGGGAAACAGGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1574  
QY 1561 gcaacaaactcccccaagggagatagggctccttggggcccttggggccatggggcaggg 1620  
Db 1575 GCAACCAATGCGCCAGGGGATGAGGTGGTCTCTGGGGCTTTCGGGACCTTGGGGCGAGGG 1634  
QY 1621 gtcgtctcctcagggcccaagcagcagcagcagcagcagcagcagcagcagcagcagc 1680  
Db 1635 GTGCTCTCTCGACAGCGCCACAGCTGTCCCTGAGAGAGAGAGGAGACAGGCGTCTGCC 1694  
QY 1681 caaacctgtccttgaaccacagcagcagcagcagcagcagcagcagcagcagcagcagc 1740  
Db 1695 CAACACTGTCTGTACCCACAGCAATTCAGAGAGAGAGCCTACAGAGGCGAGTACTCG 1754  
QY 1741 acaaaagcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1800  
Db 1755 ACAAAAGCCACAGCAGTTCAGGCTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTGCA 1814  
QY 1801 ccacaccttggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1860  
Db 1815 CCACACTGTGGCATGTGTGAGGAGAGCATCTGACCCCTGAGATTGGGAGCGCAGAGATGCC 1874  
QY 1861 cccgggaatgataataaagatactag 1887  
Db 1875 CCCGGGAATGATATAATAAGATACTAG 1901

RESULT 2  
US-09-439-313-398  
; Sequence 398; Application US/09439313.  
; Patent No. 6329505  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan Louise  
; APPLICANT: Jiang Yuqiu  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Kalos, Michael  
; APPLICANT: Fanger, Gary  
; APPLICANT: Retter, Mark

```

? APPLICANT: Solk, John
? APPLICANT: Day, Craig
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
? TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
? FILE REFERENCE: 210121.42709
? CURRENT APPLICATION NUMBER: US/09/439,313
? CURRENT FILING DATE: 1999-11-12
? NUMBER OF SEQ ID NOS: 575
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 398
? LENGTH: 278
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)..(278)
? OTHER INFORMATION: n = A,T,C or G
? OS-09-439-313-398

```

Query Match	13.9%	Score 263	DB 4	Length 278
Best Local Similarity	99.6%	Pred. No.	3.8e-48	
Matches 263	Conservative 0	Mismatches 1	Indels 0	Gaps 0

QY	981	agcagcttcgcgcacgcgcctcgcgccttcgggtggagatgtgtgcacgccacctggacatct	1040
Db	15	agcagcttcgcgcacgcgcctcgcgccttcgggtggagatgtgtgcacgccaccttgacatct	74
QY	1041	ggaagtcagcgccttcgatatgaagaagcgcactcaacctgaggcgatctcaactatgtgcct	1100
Db	75	ggaagtcagcgccttcgatatgaagaagcgcactcaacctgaggcgatctcaactatgtgcct	134
QY	1101	cgacacacctgagaggaactcgacacccgaaagcgaggtgtgacatcatggtcccgagaccca	1160
Db	135	cgacacacctgagaggaactcgacacccgaaagcgaggtgtgacatcatggtcccgagaccca	194
QY	1161	tccacctgtgagcagttctctcaaggaagtgtctactcaagccccaagctatgtgcgcttca	1220
Db	195	tccacctgtgagcagttctctcaaggaagtgtctactcaagccccaagctatgtgcgcttca	254
QY	1221	tttagtgtgctcaacaagggagaag	1244
Db	255	tttagtgtgctcaacaagggagaag	278

```

RESULT      3
US-09-213-767-1
: Sequence 1, Application US/09213767
: Patient No. 5948680
:
: GENERAL INFORMATION:
: APPLICANT: Brenda F. Baker
: APPLICANT: Lex M. Cowsett
: TITLE OF INVENTION: ANTISENSE MODULATION OF ELK-1 EXPRESSION
: FILE REFERENCE: RTS--0024
: CURRENT APPLICATION NUMBER: US/09/213,767
: CURRENT FILING DATE: 1998-12-17
: NUMBER OF SEQ. ID NOS. 47
: SEQ ID NO 1
:
: LENGTH: 2266
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (316)..(1602)
: US-09-213-767-1

```

Query Match	4.48;	Score 83.2;	DB 2;	Length 2266;
Best Local Similarity	59.88;	Pred. No. 2.5e-09;		
Matches 177; Conservative	0;	Mismatches 113;	Indels 6;	Gaps 2;

1166 ctgtgacgcttcctcaaggaagtgtctactcaagcccaacacgtatbgtccgcttcattag 1225

Db	334	ctggtgaagattctgtctgcagcgtcgtgagagag---caaggtgaatggtccaatcatctccc	390
QY	1226	tgcgtcaacaaggaagaaggcattcttcaaaatttgagagctcagcccaagtgtgcgcgtg	1285
Db	391	tgtgacttcaagggaattgtgtgtatgaattcaaaagctgtgtgtatgtcgaagagaggtgtgcgcgtg	450
QY	1286	tgtgtgcattccgaagaagccgtccgcgcgtcatgtgacatgtgacaaagcttagccgcgtcatccgc	1345
Db	451	tgtgtgacttaagcagaagaacaaagaccacaacagaaattacgaagaagctcagccgcgtcttggtg	510
QY	1346	cagctcttaacaagaagggcatcatctcggaaagccagacatctccaaagcgtctgtactaacag	1405
Db	511	tactactatgacaagaacatcatcatccgcaag---gtgtagtcggtccagaagtctgtctactaacg	567
QY	1406	tctgtgaaccccatcttgtagtgcctgtgccaaaggtccgtgaacccgcgtccaaaggggc	1461
Db	568	tctgtgtccatcaaccctgaaggtcgtcaaggtgtctcaactgtgaagatgtgcgcgtcccaagcc	623

RESULT 4  
US-08-469-412A-1  
; Sequence 1, Application US/08469412A  
; Patent No. 5856125

```

1 GENERAL INFORMATION:
2 APPLICANT: Mavrothassitis, George J.
3 APPLICANT: Blair, Donald G.
4 APPLICANT: Fisher, Ronald J.
5 APPLICANT: Real Jr., Gregory J.
6 APPLICANT: Athanasiou, Merope A.
7 APPLICANT: Sgouras, Dionysios N.
8 TITLE OF INVENTION: The ERF Genetic Locus and Its Products
9 NUMBER OF SEQUENCES: 16
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Townsend and Townsend and Crew LLP
12 STREET: Two Embarcadero Center, Eighth Floor
13 CITY: San Francisco
14 STATE: California
15 COUNTRY: USA
16 ZIP: 94111-3834
17
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: PatentIn Release #1.0, Version #1.30
23
24 CURRENT APPLICATION DATA:
25 APPLICATION NUMBER: US/08/469,412A
26 FILING DATE: 05-JUN-1995
27
28 CLASSIFICATION: 435
29
30 ATTORNEY/AGENT INFORMATION:
31 NAME: Garrett-Wackowski, Eugenia
32 REGISTRATION NUMBER: 37,330
33 REFERENCE/DOCKET NUMBER: 015280-229000
34
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: (415) 576-0200
37 TELEFAX: (415) 576-0300
38
39 INFORMATION FOR SEQ ID NO: 1:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 2667 base pairs
42 TYPE: nucleic acid
43 STRANDEDNESS: single
44 TOPOLOGY: linear
45
46 MOLECULE TYPE: cDNA
47
48 FEATURE:
49 NAME/KEY: CDS
50 LOCATION: 123..1769
51 OTHER INFORMATION: /note="human ERF (ETS2 Repressor
52 OTHER INFORMATION: Factor) cDNA"
53
54 US-08-469-412A-1

```

Query Match	4.08;	Score	75.2;	DB	2;	Length	2667;
Best Local Similarity	59.18;	Pred. No.	1.3e-07;				
Matches	150;	Conservative	0;	Mismatches	98;	Indels	6;
						Gaps	1;







```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Weiser & Associates
STREET: 230 South Fifteenth Street
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: AEDIT 1.0 DOS text editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343.443B
FILING DATE: 18-NOV-1994
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00494
FILING DATE: 19-MAY-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: FR 92/06123
FILING DATE: 20-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989.6121P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2938 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

FEATURE:
NAME/KEY: CDS
LOCATION: 143..1498
US-08-343-443B-3

Query Match
Best Local Similarity 3.7%; Score 71; DB 2; Length 2938;
Matches 132; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 1156 gccatccacacctgagcagctcccaaggagtgctactcaagccccaagctatgagccg 1215
DB 979 gcagatccagcctgtggcgaattctcctggagcgtctccgcacagcccaagccagctg 1038
QY 1216 ctctaatgagtgctcaacaagagagagcagctcctcaaatgagagctcagccagct 1275
DB 1039 tatcaccctggagggcagccacacggcgag-----ttcaaaatgacggaccccgatgagct 1092
QY 1276 ggcgcggcgtggtggagcgcgaagacgctccgcatgaaactggaagaactgagccg 1335
DB 1093 ggccagagcctggggcagcggaaagcaagcccaatgaaatttcgacaaagctgagccg 1152
QY 1336 ctccatccgcagctctacaagaaggagcagctat 1368
DB 1153 ggccctccggttatctatgattaaaacatttat 1185

```

```

STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210-4433

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/780.835B
FILING DATE: 10-JAN-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: UTSC500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2410 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-780-835B-1

Query Match
Best Local Similarity 3.7%; Score 70.2; DB 2; Length 2410;
Matches 158; Conservative 0; Mismatches 103; Indels 2; Gaps 2;

QY 1109 gagagagctgagccgaagagagtgactcaccatcagctccgagcccatccactg 1168
DB 1308 gaaggatgtgagctttccggaggagccacccttaccagccgcgggtgctttacaactg 1367
QY 1169 tgcagctccctcaaggagtgctactcaagccccaagctatgagcgtcattatgagtg 1228
DB 1368 tggcaggttctc---cgtggccctctgagatgacccacaaatgctcattatgcttgg 1424
QY 1229 ctcaacaaggagagagcagctctcaaatgagagctcagccagctgagccgctg 1288
DB 1425 acagcccgccggaatgc---actttaactaatgtgaacctgaagcaggttgcagcctcgg 1481
QY 1289 ggcacccgcaagacccgtccgcgcctaatcagctgacagcagctgagcgcctcatcgcag 1348
DB 1482 ggtatccagaaacggccagccatgattatgacaaagctgagccgctcgtccgattac 1541
QY 1349 tcttacaagaaggagcagctatccggaag 1375
DB 1542 tatatatgaaaggaatcagcagacag 1568

RESULT 11
US-09-303-268-1
Sequence 1, Application US/09303268
Patent No. 6172212
GENERAL INFORMATION:
APPLICANT: Hung, Mien-Chie
TITLE OF INVENTION: Xing, Xiangming
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,268
; FILING DATE: 30-Apr-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/780,835
; FILING DATE: 10-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: UTSC500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
;
; INFORMATION FOR SEQ ID NO: 1:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2410 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-303-268-1

```

```

Query Match          3.7%; Score 70.2; DB 4; Length 2410;
Best Local Similarity 59.2%; Pred. No. 1.6e-06;
Matches 158; Conservative 0; Mismatches 103; Indels 6; Gaps 2;

```

```

QY 1109 gaggaagactgacgcagacgaagtgtgactcattcgtccggcgagccatccactg 1168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1308 GAAGGATTGAGCTTCCGCGGAGCGCCACCTACCGACCGCGGCTGCTTACAACTG 1367
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1169 tggcagttctcaagaagttgctcaagccccaagctatggccgcttaagttg 1228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1368 tggcagttctc---gctgacctgctgctgacccacacaaatgctcatttattgcttg 1424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1229 ctcaacaaggaagagcactctcaaaattgagactcagcccgagtgccggctgtg 1288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1425 ACAGCGCGGGAATG---AGTTAACTAATTGAACTGAAGAGGTGCGAGGCTCTGG 1481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1289 ggcattccgaagaacgcgtccgcatgacgtcgcaagctcgagccgtccatccgcaag 1348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1482 GGTATCCAGAAACCGGCGCATGATATGACAAGCTGAGCCGCTCGCTCGATAC 1541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1349 tcttacaagaaggcattcccggaag 1375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1542 TATTATGAGAAAGCATCATGCAAGAAG 1568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 12
US-09-116-049-1
; Sequence 1, Application US/09116049A
; Patent No. 6248351
; GENERAL INFORMATION:
; APPLICANT: Hung, Men-Chie
; TITLE OF INVENTION: HUMAN PEA3 IS A TUMOR SUPPRESSOR FOR CANCER CELLS
; FILE REFERENCE: UTSC:582
; CURRENT APPLICATION NUMBER: US/09/116,049A
; FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2410
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-116-049-1

```

```

Query Match          3.7%; Score 70.2; DB 4; Length 2410;
Best Local Similarity 59.2%; Pred. No. 1.6e-06;

```

```

Matches 158; Conservative 0; Mismatches 103; Indels 6; Gaps 2;
QY 1109 gaggaagactgacgcagacgaagtgtgactcattcgtccggcgagccatccactg 1168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1308 GAAGGATTGAGCTTCCGCGGAGCGCCACCTACCGACCGCGGCTGCTTACAACTG 1367
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1169 tggcagttctcaagaagttgctcaagccccaagctatggccgcttaagttg 1228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1368 tggcagttctc---gctgacctgctgctgacccacacaaatgctcatttattgcttg 1424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1229 ctcaacaaggaagagcactctcaaaattgagactcagcccgagtgccggctgtg 1288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1425 acagcgcgggagatg---agtttaactaattgacctgaagaggtgccaagctctg 1481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1289 ggcattccgaagaacgcgtccgcatgacgtcgcaagctcgagccgtccatcccgcaag 1348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1482 GGTATCCAGAAACCGGCGCATGATATGACAAGCTGAGCCGCTCGCTCGATAC 1541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1349 tcttacaagaaggcattcccggaag 1375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1542 TATTATGAGAAAGCATCATGCAAGAAG 1568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 13
US-09-360-779-1
; Sequence 1, Application US/09360779
; Patent No. 6268216
; GENERAL INFORMATION:
; APPLICANT: Denneris, Bryan S.
; APPLICANT: Rydoro, Dmitry V.
; APPLICANT: Hendricks, Timothy J.
; TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
; FILE REFERENCE: Case-03828
; CURRENT APPLICATION NUMBER: US/09/360,779
; CURRENT FILING DATE: 1999-07-26
; EARLIER APPLICATION NUMBER: 60/094,264
; EARLIER FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (112)..(1131)
US-09-360-779-1

```

```

Query Match          3.4%; Score 64.8; DB 4; Length 1752;
Best Local Similarity 53.1%; Pred. No. 2.1e-05;
Matches 186; Conservative 0; Mismatches 137; Indels 7; Gaps 2;

```

```

QY 1156 gccatccactgtgacggttctcaagaagttgctactcaagccccaagctatggcg 1215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 555 gaagatccagttgtgcaagtttctactgagctgtgcaagccgcgaagcgcgctg 614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1216 cttaattagttgtctcaacaaggaagagcatcttcaaaattgagactcagcccggt 1275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 615 catcgctgggagggc-----gccaagcgaggttcaagctcagcaccgcgaggt 668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1276 ggcgcgctgtgtggcattccgcaagaacggtccgcgaactgtgcaagaagctgagcg 1335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 669 ggcgcgagctgtggcgagcgcaagaagcaccataatgactaagcaagctaaagctg 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1336 ctccatccgcagttcttcaagaagagcatatccggaagccaagatctccagcgct 1395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 729 agcaatgcgtactactaagaaaaaacatcatagcaagtgtagcggaagcgctagc 788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1396 cgtctacagttgtgcaccccatctgagctgtgcccgaaggtgaaacccgcctca 1455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 789 ctaccgcttgcttccagagcgctgtgcaagcgttgcacgacacacacccgcgacgc 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 18:16:11 ; Search time 3492.95 Seconds

(without alignments)  
11347.102 Million cell updates/sec

Title: US-09-841-963A-1

Perfect score: 1894

Sequence: 1 gctgactctctccagcac.....ataaagatactagagactg 1894

Scoring table: GAPOP 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :

1: gb\_ba:\*  
2: gb\_hcy:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hcy\_hum:\*  
31: em\_hcy\_inv:\*  
32: em\_hcy\_other:\*  
33: em\_hcy\_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No. Query Score Match length DB ID Description

Result No.	Query	Score	Match length	DB	ID	Description
1	AX048148	1870	98.7	1894	6	AX048148 Sequence
2	1870	98.7	1894	9	AF071538	AF071538 Homo sapi
3	1868.4	98.6	1918	6	AX175516	AX175516 Sequence
4	1865.2	98.5	3317	6	AX048159	AX048159 Sequence
5	1864.4	98.4	1942	9	BC021299	BC021299 Homo sapi
6	1811.8	95.7	1898	9	AB031549	AB031549 Homo sapi
7	757	40.0	1856	10	BC012648	BC012648 Mus muscu
8	738.4	39.0	1704	10	AB019436	AB019436 Mus muscu
9	643	33.9	150660	9	AL157372	AL157372 Human DNA
10	329.8	17.4	1751	6	AX048155	AX048155 Sequence
11	263	13.9	278	6	AX106617	AX106617 Sequence
12	263	13.9	278	6	AX140908	AX140908 Sequence
13	263	13.9	278	6	AX200768	AX200768 Sequence
14	263	13.9	278	6	AX267424	AX267424 Sequence
15	156.6	8.3	105772	2	AC095858	AC095858 Rattus no
16	151.2	8.0	105772	2	AC095858	AC095858 Rattus no
17	147.6	7.8	362	3	DRODTSAA	DRODTSAA Droso
18	147.6	7.8	38178	2	AC015374	AC015374 Droso
19	147.6	7.8	180887	3	AC008253	AC008253 Droso
20	147.6	7.8	188549	3	AC007818	AC007818 Droso
21	147.6	7.8	193245	3	AC008217	AC008217 Droso
22	147.6	7.8	230940	3	AE003762	AE003762 Droso
23	144.4	7.6	1832	3	AF106964	AF106964 Strongylo
24	96	5.1	5315	3	DROE74B	DROE74B Droso
25	96	5.1	6479	3	DROE74A	DROE74A Droso
26	96	5.1	51405	2	AC019512	AC019512 Droso
27	96	5.1	174643	3	AC009371	AC009371 Droso
28	94.4	5.0	4800	3	DM74E	DM74E Droso
29	87.2	4.6	148462	2	AC095263	AC095263 Rattus no
30	83.2	4.4	2266	6	AR073028	AR073028 Sequence
31	83.2	4.4	2266	9	HOMEK1A	HOMEK1A Homo sapi
32	83.2	4.4	2266	9	AB016193	AB016193 Homo sapi
33	81.6	4.3	848	5	AF075706	AF075706 Gallus ga
34	81.4	4.3	125020	9	AF429315	AF429315 Gallus ga
35	81.2	4.3	742	9	AF000672	AF000672 Homo sapi
36	80.8	4.3	1436	5	DRE249590	DRE249590 Danio rer
37	80.4	4.2	660	10	MMELK1P	MMELK1P Mus muscu
38	80.4	4.2	2286	6	AX305989	AX305989 Sequence
39	80.4	4.2	2286	10	MMELK1GEN	MMELK1GEN Mus muscu
40	79.6	4.2	2508	5	AF075708	AF075708 Gallus ga
41	79.2	4.2	4200	5	DRU84615	DRU84615 Danio rerio
42	78.8	4.2	742	9	HSN339670	HSN339670 Homo sapi
43	78.8	4.2	1403	9	HSDNAFEV3	HSDNAFEV3 Homo sapi
44	78.8	4.2	1901	9	HSRNF4EV	HSRNF4EV Homo sapi
45	78.8	4.2	169741	9	AC097468	AC097468 Homo sapi

## ALIGNMENTS

Result No.	Query	Score	Match length	DB	ID	Description
1	AX048148	1894 bp	DNA	11near	PAT 15-DEC-2000	
LOCUS	AX048148	Sequence 1 from Patent WO0070092.				
DEFINITION	AX048148					
ACCESSION	AX048148					
VERSION	AX048148.1	GI:11876967				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
Source						
CDS						

protein\_id="CAC19002.1"  
/db\_xref="GI:11876968"  
/translation="MGSAAGSLGSEVSPSHLLLPDVTYRRTGLEKAAGAVGLEERDMS  
PSPAPEQGI SAFLVLYSPMKLYPEDSSMAKAGASREPEPCQVDTSDCAQD  
GSDLYRGCGLTLEHSTLEQWQSWYGVGLIDTFACTLNTADPMQMSPEWQKILL  
WTEQYRLPMPGKAFQELAGLELCAKESEEDPROSPILGCVLYLHLLHDIKRSAAAMKEK  
TSPRAHICASTSESWTDESVDSGCGPFLHMQFLLELLKHAHSTGRITRLNKEK  
GIFLEISDAQVARKLWGIRKNRPNANNRYKLSRYKNGILRRPDISQRLVQYVFH  
I"

BASE COUNT	368	a	653	c	571	g	302	t
------------	-----	---	-----	---	-----	---	-----	---

[illegible]

QY	841	catcacgcagatcccatctgaacctcggagccccaagctgacgaagctgcccctgtgac	900
Db	841	CATACCGCAGATCCCATGCACTGGAGCCCCAGCAATGTGCAGAAAGTGGCTCTGTGAC	900
QY	901	agaagcaccaataccggtctgcccccatctgggcaagcctctccagaagcttgcgcgcaaga	960
Db	901	AGAGCACCAATACCGGCTGCCCCCATGGGCAAGGCGCTTCCAGGAGCTGGCGGCAAGA	960
QY	961	gctgtgcacatgttcggaagaagaagttccgcgaagcctcgcgccctgggtgggtgtgtgtc	1020
Db	961	GCTGTGGCCATGTTCGGAGGAGAGAGTTCCGCCACAGCGCTTGGCCCTGGGTGGGAGTGTCT	1020
QY	1021	gcacgcccaccttgaacatccgaagatcgaagccttgatgaagaagcagacttcaactctg	1080
Db	1021	GCACGCCACCTGGACACTCTGGAAGTCAGGGGCGTSTGGATGAAGAAGCGGACTTCACCTCG	1080
QY	1081	ggcgatctcaactatgtgtcctcgaacagtgagaagcttgacgcagcagcgagttgactc	1140
Db	1081	GGCGATTCTACTACTGTGTGCTTCGCACAGATGGAGAGACTTGACCGACGAGCGGTGGATCTC	1140
QY	1141	atcatgtctccgggagagcccatccacctgtggcagattctctccaaaggagttgtctactcaaac	1200
Db	1141	ATCATGTCTCCGGGAGGCCCATCCACTGTGGAGTTCTCTCAAGGAGTTGCTACTCAAGCC	1200
QY	1201	ccacagctatgcccgtctctcatagtgtgtcacaacaaagaaaggcattctcaaatgtga	1260
Db	1201	CCACAGCTATGGCGGCTTCAATTAGTGGGCTCAACAAGGAGGAGGCACTTCAAAATATGA	1260
QY	1261	ggactctcagcccaggttgcccgcgctgtgtgggacatccgcaagaaacgctccgcgcatagaact	1320
Db	1261	GGACTCACCCCAAGGTGGCCCCGCGCTGTGTGGGCATCCGCAAGAACCGTCCCGCCATGAACCTA	1320
QY	1321	cgaacagcttgcgcgcctccatccgcagctcttcaagaagaaggacatccgcgaagcaga	1380
Db	1321	CGACAGCTGAGCGCGCTCATTCGCCAAGTTTACAGAAGAGGCGATCATCCGGAAGCCAGA	1380
QY	1381	catctccagcgcctctgtctcaagttcgttgcacccacatgtgagtgcctgagccagggcc	1440
Db	1381	CATCTCCAGCGCCTCTGTCTTACCAAGTTCGGCACCCCACTCAATGATGCTGGGCCAGGGCC	1440
QY	1441	tgaaaccccgcccttaaggggcctctctctcgtccgcgcctcctccagcagggcccttgaaga	1500
Db	1441	TGAACCCCGCCCTAAGGGGCGCTCTCTCTCTGCTCGCCCTGCTCAGCCAGGCCCTGAGATG	1500
QY	1501	ggggaaaaagggcaggtgtgtctctgtgtctcttgacctctccagaagcccaaggtccaaggag	1560
Db	1501	GGGGAAAAAGGGGAGTGTGCTGTGCTGCTCTGACCTTCCAGAGCCCAAGTCAAGGAGG	1560
QY	1561	gcaacccaactcgcaccaagggggataatggtctctcttgggctctcgggacatggggcagag	1620
Db	1561	GCAACCAACTGCCCAAGGGGATATAGGTCTCTTGGGGGCTTCGGGACATGGGGCAAGG	1620
QY	1621	gtgtctcctctcgaagcccaagctgtctccccttgaaagaaagaaaggaacagggctctccc	1680
Db	1621	GTGTCTCTCTCTCAAGCCCAAGCTGTCTCCCTGGAGGACAGAGGAGACAGCGCTCTCTCC	1680
QY	1681	caacacctgtcctctgaccccaagatctccagagcagaagcctcaagaaaggcagtgactcg	1740
Db	1681	CAACACCTGTCTTGACCCCAAGCATTTCCAGAGCAAGGCTCTACAAAGAGGCGAGTGACTCG	1740
QY	1741	acaagaagcacaagcagttcgaagcctctctctgtctcatcccccctgcgctcccatctctga	1800
Db	1741	ACAAAGGCCACAGGCAAGTCCAGGCGTCTCTCTCTCTCCATCCCCCTGCTCCCATTTCTCA	1800
QY	1801	ccacacctcgcagatgtgcagggagaacatctgcacccctcagttctgggcagccagaaatgac	1860
Db	1801	CCACACCTGCGCATGTGTCAGGAGGAGACATCTGCACCCCTGATGTGGGCAAGCCGAGAGTGCC	1860
QY	1861	cccggaatgagataaagaatactagaaactg	1894
Db	1861	CCCGGAATGAGATAAAGATTAAGTTACTAAGAACTG	1894



RESULT 2  
AF071538 1894 bp mRNA linear PRI 10-JUN-2000  
LOCUS Homo sapiens Ets transcription factor PDEF (PDEF) mRNA, complete  
DEFINITION cds.  
ACCESSION AF071538  
VERSION AF071538.1 GI:4007417  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1894)  
AUTHORS Oettingen,P., Finger,E., Sun,Z., Akbarali,Y., Thamrongsak,U.,  
Boltax,J., Grall,F., Dube,A., Weiss,L., Brown,L., Quinn,G., Kas,K.,  
Endress,G., Kunsch,C. and Libermann,T.A.  
PDEF, a novel prostate epithelium-specific ets transcription  
factor, interacts with the androgen receptor and activates  
prostate-specific antigen gene expression  
J. Biol. Chem. 275 (2), 1216-1225 (2000)  
2 (bases 1 to 1894)  
Libermann,T.A., Oettingen,P., Kunsch,C. and Finger,E.  
Direct Submission  
Submitted (10-JUN-1998) Medicine, Beth Israel Deaconess Medical  
Center, 330 Brookline Ave., Boston, MA 02215, USA  
LOCATION/Qualifiers  
1. 1894  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="6p21.3"  
/tissue-type="prostate"  
1. 1894  
/gene="PDEF"  
416..1423  
/note="PDEF"  
/note="prostate epithelium-specific"  
/codon\_start=1  
/product="Ets transcription factor PDEF"  
/protein\_id="AAC95296.1"  
/db\_xref="GI:4007418"  
/translation="MGSASPLSSVSPSHLLPDDVYSRITGLEKAAAGAVLERRDMSPSPAPPEGLISAFVLSYFDMLYPEDSSNAKAPKASSREPERPEECVIDSOAPA  
GSDLVPRGGLTLEHSEHQYQSMYGEVLKDIETACKLNTITADPMDSPSNVQKML  
WTEQITLPPMGKAFQELAKKEJCAHSEDFRORSPLGGVLAHLIDIKLSAMKKEK  
TSPGAIHYCASTSESWTSDVSDSCSGQIHLMOFLKEDLKPHSYGRFIRMLNKEK  
GIFKIEDSAQVARIWMGIRKNRPAMNYDKLSIRIROYKKGIIRKPDISORLYOFEVHP  
1"

BASE COUNT 368 a 653 c 571 g 302 t  
ORIGIN

Query Match 98.7%; Score 1870; DB 9; Length 1894;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1879; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 gcttactctcccaagacatctctgactctgcccgtgtcacaacagcccccaagacc 60  
DB 1 gttctactctcccaagacatctctgactctgcccgtgtcacaacagcccccaagacc 60  
QY 61 agtccccaagcctgtcgcagctccctgcagccctcaggttggtgacctgcacagtg 120  
DB 61 agtccccaagcctgtcgcagctccctgcagccctcaggttggtgacctgcacagtg 120  
QY 121 ccagagcagcagcctggcctggcctggcctggcctggcctggcctggcctggcctgg 180  
DB 121 ccagagcagcagcctggcctggcctggcctggcctggcctggcctggcctggcctgg 180  
QY 181 gccac 240  
DB 181 tcagagggcgac 240

QY 241 gccac 300  
DB 241 gccac 300  
QY 301 cagtgacctcagcttgccacacacacacacacacacacacacacacacacacacac 360  
DB 301 cagtgacctcagcttgccacacacacacacacacacacacacacacacacacacac 360  
QY 361 ctccttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
DB 361 ctccttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
QY 421 cagcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480  
DB 421 cagcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480  
QY 481 ggttcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 540  
DB 481 ggttcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 540  
QY 541 ctggagctcccaagctcccaagcagcagcagcagcagcagcagcagcagcagcagc 600  
DB 541 ctggagctcccaagctcccaagcagcagcagcagcagcagcagcagcagcagcagc 600  
QY 601 ctgtacacatgctgtac 660  
DB 601 ctgtacacatgctgtac 660  
QY 661 ttggagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
DB 661 ttggagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
QY 721 gggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780  
DB 721 gggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780  
QY 781 gcatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840  
DB 781 gcatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840  
QY 841 catcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900  
DB 841 catcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900  
QY 901 aagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960  
DB 901 aagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960  
QY 961 gcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1020  
DB 961 gcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1020  
QY 1021 gacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080  
DB 1021 gacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080  
QY 1081 ggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140  
DB 1081 ggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140  
QY 1141 atcatgctccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200  
DB 1141 atcatgctccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200  
QY 1201 ccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260  
DB 1201 ccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260  
QY 1261 ggaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1320  
DB 1261 ggaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1320  
QY 1321 cgaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380

```

|||||
Db 1321 CGAAGAGTGAAGCGGCTCCATCCGCAATATACAAAGAGGCAATCATCGGAAGCCAGA 1380
|||
Qy 1381 catctccagcgcctcgtctacagctcgtgcaaccacatctagtgcctgcccagacc 1440
|||
Db 1381 CATCTCCAGCGCGCTGCTGTACAGTTCGTGCAACCCATCTAGTCCCTGGCCAGGGCC 1440
|||
Qy 1441 tgaaccgcgcctcagggcctctctcctcgtcctcgtcctcctcagcagggcctagatc 1500
|||
Db 1441 TGAACCCCGGCTCAGAGGCGCTCTCTCTGCTGACCCCTCCTCAGCCAGGCCCTGAGATG 1500
|||
Qy 1501 ggggaaacagcagctgctcgtcctcgtcctcgtcctcgtcctcgtcctcgtcctcgtc 1560
|||
Db 1501 GGGGAAACAGGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
|||
Qy 1561 gcaacacacgtccagcaggggagatagtgctcctcgtcctcgtcctcgtcctcgtcctc 1620
|||
Db 1561 GCAACCAACTGCCCCAGGGGGATATGGTCTCTGGGGCTTCCGGACCATGGGGCAGGG 1620
|||
Qy 1621 gttcttctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1680
|||
Db 1621 GTCTTCTCTCTCAGGCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
|||
Qy 1681 caaacacctgctcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1740
|||
Db 1681 CAACACCTGCTGCTGAGCCAGCATTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
|||
Qy 1741 acaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1800
|||
Db 1741 ACAAAAGCCACAGGCAAGTCCAGGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
|||
Qy 1801 ccaaacctgagcagctgctcagcagcagcagcagcagcagcagcagcagcagcagcagc 1860
|||
Db 1801 CCACACCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
|||
Qy 1861 cccggagatgagataaagaatactagagagactg 1894
|||
Db 1861 CCCGGGAATGATTAATAAAGTACTAGAGAACTG 1894
|||
RESULT 3
AX175516 1918 bp DNA linear PAT 03-JUL-2001
LOCUS AX175516
DEFINITION Sequence 1 from Patent WO0142472.
ACCESSION AX175516
VERSION AX175516.1 GI:14598855
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1918)
Harkins, R., Lin, R.J., Luke, M., Montecarlo, F., Parkes, D., Parry, G.,
Steinbrecher, R., van Heuvel, P.T. and Xuan, J.A.
dna encoding a novel prostatic polypeptide
Patent: WO 0142472-A 1 14-JUN-2001;
TITILE SCHERING AKTENSELUSCHAFT (DE)
JOURNAL Location/Qualifiers
FEATURES
source
1. 1918
/organism="Homo sapiens"
/db_xref="taxon:9606"
422..1429
/note="unnamed protein product"
CDS
/codon_start=1
/protein_id="CAC43877.1"
/db_xref="GI:14598855"
/translation="MGSASPGLSVSPSHLLPDPVSRGTLEKAAAGVGLERDWS
PSPRPBGLSALFYLSYDMLXPEDSSMAAKRPAASSSEPEPEPDCPVDSQAPA
GSLDLPGLLTLEBSLEVOVMYGEVKDLETCKLINTADPMDSFVQKML
WREHOYLRPPMKAFORDELAKELCAMSEDFORRPLGDVJHAHLDTKSSAMWKEK
TSPGAIHCASTSEBSWDSSEVSSCGSPHLIOLPFLKELKPSYSORFRLMKKEK
GIFKIDSQAVARLMGIRKRNRPAMMYDKLSRSIROYYKKGIIRKPDISORLIVYQFVHP
I"

```

```

BASE COUNT 385 a 656 c 574 g 303 t
ORIGIN
Query Match 98.6%; Score 1868.4; DB 6; Length 1918;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 1 gttcgtactctcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 60
|||
Db 7 gttgactctctcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 66
|||
Qy 61 agtcctcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 120
|||
Db 67 agtctcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 126
|||
Qy 121 ccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 180
|||
Db 127 CCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 186
|||
Qy 181 gcaacacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240
|||
Db 187 TCAGAGAGGCGCACCCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
|||
Qy 241 gcaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 300
|||
Db 247 GCCACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 306
|||
Qy 301 cagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 360
|||
Db 307 CAGTGGCTCAGCTGAGCCACACCTCTTCCGGCCCTGAAAGTGGCACTCAGCAGACAG 366
|||
Qy 361 ctccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420
|||
Db 367 CTCCTGGGACACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 426
|||
Qy 421 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 480
|||
Db 427 CAGCCGCCAGCCCGGGGTGAGACAGCAGCTATCCCCAGCAGCAGCAGCAGCAGCAGCAGC 486
|||
Qy 481 ggtgtcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540
|||
Db 487 GGTGTGCGGAGACAGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
|||
Qy 541 ctgagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 600
|||
Db 547 CTGAGATCCAGTCACAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCT 606
|||
Qy 601 ctttgacatgctgtacacctgagcagcagcagcagcagcagcagcagcagcagcagcagc 660
|||
Db 607 CTTTGACATGCTGTATCCCTTGAGAGACAGCAGCTGGGACGCAAGGCCCTTGAGGCGCAG 666
|||
Qy 661 tcggagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
|||
Db 667 TCGGAGAGAGAGCAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726
|||
Qy 721 gggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 780
|||
Db 727 GGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 786
|||
Qy 781 gcaatcagctgtgtgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
|||
Db 787 GCAATCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 846
|||
Qy 841 catcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 900
|||
Db 847 CATCACCCAGATCCCATGAGACTGAGAGCCAGCAAGTGTGAGAGAGAGAGAGAGAGAGAG 906
|||
Qy 901 agagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 960
|||
Db 907 AGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 966
|||
Qy 961 gctgtgcgcagctgtcggagagcagcttcgcagcagcagcagcagcagcagcagcagcag 1020

```

Db	967	GCCTGTCGCGCATGTCGAGAGAGATGTCGCGAGCCCTGCCCCGGCGGGAGATGCT	1026
Qy	1021	gcaaccccaacttgagacatctggaagtcagcggcccttgatgaagaagcagactcaccctgg	1080
Db	1027	GCACGCCCACTGACATCTGGAAGTCACGGCCCTGATGAAAGACGGACTTTCACCTGG	1086
Qy	1081	gagcatcactactgtgctcgcaccagtagagagagctgagaccgaacgaggtgagctc	1140
Db	1087	GGGATTACTACTGTGCTCGACAGTAGAGAGACTGGACCGACAGAGAGTGGACTC	1146
Qy	1141	atcatgtctccggcgagcccaatcactctgtggaagttcctccaaggaggttgctactaagcc	1200
Db	1147	ATCATCTCCGGGACACCCATCCACTGAGTTCCTCAAGAGAGTTCTACTCAAGCC	1206
Qy	1201	ccacacataatgagcctcattagatgagtcacaaagagagagagccatctcaaaatiga	1260
Db	1207	CCACACTATGCGCCCTTATTAAGTGGCTCAACAAGAGAGGCACTTTAAATTTGA	1266
Qy	1261	ggaactaagccaggtggccggcctgtggggcattccgcaagaacgctcccgcaatgaacty	1320
Db	1267	GGACTAGCCAGAGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCGCCATGACTA	1326
Qy	1321	cgagcaagctgagcgcgtccatccgcagctctcaagaagagggatcatccggaagccaga	1380
Db	1327	CGCAAGCTGAGCCGCTCATCGCCAGATTAACAAGAGGGGCACTATCCGGAAGCCACA	1386
Qy	1381	catctccagcgcctcgtctacacagttgtagcccccacttgatgacctgagccagggcc	1440
Db	1387	CATCTCCAGCGCTGCTGTACAGTTCTGACACCCCATCTGATGCTGCGCCAGGGCC	1446
Qy	1441	tgaacccgcgcctcaggggcctctctcgtcgcctcgcctcgcctcgcctcgcctcgcctc	1500
Db	1447	TGAACCCCGCCCTCAAGGGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1506
Qy	1501	ggggaaagcggagtgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	1560
Db	1507	GGGAAAGCGGCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1566
Qy	1561	gcaacccaactgcccccaaggagatagtgctcctcgtcgcctcgcctcgcctcgcctcgcctc	1620
Db	1567	GCAACCAACTGCCCGCAAGGGGAGATAGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1626
Qy	1621	gtgctcctcctcagggcccaagctgctcctccctcggagagagagagagagagagagagag	1680
Db	1627	GTCTCTCTCTCTCAGGCGCCAGCTGCTCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAG	1686
Qy	1681	caacacactgctctgaccccaagatttcagagagagagagagagagagagagagagagagag	1740
Db	1687	CAACACTCTGCTCTGACCCCAAGATTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1746
Qy	1741	acaaagggcccaagcagctcagggcctctctcgtcctcctcctcctcctcctcctcctcctc	1800
Db	1747	ACAAAGGCCAGAGAGTCCAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1806
Qy	1801	ccacacactgagcagtggtcagggagagacatctgacccctcaggttgagcagcagagagag	1860
Db	1807	CCACACTCTGCGATGCTGAGGAGAGAGATCTGACCCCTGAGTTGGGCAAGAGAGAGAG	1866
Qy	1861	cccgggaatgatacaagatactagagaacty	1894
Db	1867	CCCGGGAATGATTAATTAAGATTAAGAGACTG	1900

RESULT 4  
LOCUS AX048159 3317 bp DNA linear PAT 15-DEC-2000  
DEFINITION Sequence 12 from Patent WO0070092.  
ACCESSION AX048159  
VERSION AX048159.1 GI:11876975  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 3317)				
TITLE	Kaufmann, J. and Xin, H.				
JOURNAL	Expression of ets-domain proteins in cancer				
	Patent: WO 0070092-A 12 23-NOV-2000;				
	Chiron Corporation (US)				
FEATURES	Location/Qualifiers				
source	1..3317				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/note="GSEF-encoding sequence with promoter (Figs. 2A-2B)"				
	1..1380				
	/note="Promoter"				
	1815..2822				
	/note="GSEF coding sequence"				
	/codon_start=1				
	/protein_id="CAC19004.1"				
	/db_xref="GI:11876976"				
	/translation="MGSASPLGLSSVPSHLLPDPVSRHTGLEKAAAGAVGERDWS PSPATPEGLSAFLSYFDMLEPDSMAKAPGASREPEPEECVIDSOAPA GSDLVVPGGLIEHLSLEQVSNVGEVLEKIDETACKLINTADPMDSPSVOKRL WTEHOYRLPMKARFOELAKELCAASEQFORPSPLGSDVTHAHLDIKSAAMKER TSPGAIHYLCASSTSESWTDSVDSGQPLHMQPLKELLPKPSYGRFIRLNKEK GIFIEDSAQVARKLWGIIRKRPAMNYDKLSRSIRQYKKGIIRKPDISQRLVYQFVHP I"				
BASE COUNT	710 a	1026 c	970 g	611 t	
ORIGIN					
Query Match	98.5%; Score 1865.2; DB 6; Length 3317;				
Best Local Similarity	99.0%; Pred. No. 0;				
Matches 1876; Conservative	0; Mismatches 18; Indels 0; Gaps 0;				
Qy	1 gctgactcctcccaagacattcctgactcgtcgtcgcacacgcccccaagacc	60			
Db	1400 gctgactcctcccaagacattcctgactcgtcgtcgcacacgcccccaagacc	1459			
Qy	61 agtctccaaagcctgctgcccagctcctcgcagagccctcagagtggtgctcgcagagt	120			
Db	1460 AGTCTCCAAAGCCTGCTGCGCAGCTCCTGCAAGCCCTCAGAGTTGGGCTTGCCAGCGTG	1519			
Qy	121 ccagagagcagccttggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt	180			
Db	1520 CCAGAGGACAGCCCTGGGCTGGGGTGAAGGACATCCCTACAGCAGCAGCCTGAGACCC	1579			
Qy	181 gccacacagccaccccttgaagggttgccagagcccaagtgccaaactgagtgctgctct	240			
Db	1580 TCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAAGTGGCAACTGAGTGTCTCTCT	1639			
Qy	241 gccacacagcccttgcgtcccttggttccgctggtgcccccaagatgctgctgtagacagc	300			
Db	1640 GCCACAGCCCTGCTGGGCGCTGGTCCGGTGGCCGCCCAAGATCTGCGGTGAGACAGCG	1699			
Qy	301 cagtgagctcagctggtccacacccctcctccgggccccctggaggttgagcagtcagcagag	360			
Db	1700 CAGTGGCTCTAGCTGCGCCACACCTTCCCGGCCCTGAAGTTGGCACTCAGCAGACAG	1759			
Qy	361 ctccctgggcaacagcagctaaacagacagcccgccagcccaacagcagcagcagcagcag	420			
Db	1760 CTCCTGGGCAACAGCAGCTAACAGACACAGCCGCCCAACAGCAGCGGCAAGGGCATGGG	1819			
Qy	421 cagcgccagcccggtctcagacagcgatcccccaagcaccctcgtcgtcccccagac	480			
Db	1820 CAGCGCCAGCCCGGCTGAGACAGCTATCCCAAGCACCTCTGCTGCCCCCGGACAC	1879			
Qy	481 ggtgtcgagagcagcttgagagagagcgagcaggggaggttggttctgagagagagga	540			
Db	1880 GGTGTGCGGAGAGCGCTTGAGAGAGCGGCAAGCGGGGAGTGGGTTCAGAGACGGGA	1939			
Qy	541 ctgagagtcacagtcacacccgcaagcccgagcagggcctgctcagcctcagcctcctccta	600			
Db	1940 CTGAGAGTCCAGTCCACCCGCGACCGGAGCAGAGGCGCTGCTGCTTCACTCTCTCA	1999			



```

"protein_id"="AA021299.1"
/contig="GI:18204266"
/translation="MGSAAPGLSSVSPSHLLLPDVTYRFTGLEKAAAGLGERDMS
PSPATPEEGISAFYILSYFDMLPEDESSMAKAGASSREPEPEPCQVTSQAA
GSLDIPGLGLLEHSHLEQVOSMYAGVELDIEACKILNTAPADMSPSSYQKILL
WTEHQYRLPMPGKAFOELAGCELCAMSEEDGROSPGLGVLNHAHLIDIMKSAAMKER
TSPCAIHVCASTSESWTDSFVDSGGOPIHMQFLKELLIKHPSYGRFLPWLNKKR
GIFRLEDSAVYARLMGIRKNRPANNYTKLSRSTQYTKKGIIRKPDLSQALVYQFVHP
I"
BASE COUNT      411 a      656 c      575 g      300 t
ORIGIN

```

Query Match	98.4%	Score 1864.4	DB 9	Length 1942
Best Local Similarity	99.2%	Pred. No. 0		
Matches 1874	Conservative	0	Mismatches	16
			Indels	0
			Gaps	0
QY	5 gactccctcccccagacatctctgacactctgctgtgtccacactgcccacagaccagtc	64		
Db	10 GACTCTCTCTCCAGCAGCAATTCCTGCACTCTGGCGGTGTCCACACTGCCCCAGACCCAGTC	69		
QY	65 ctccaaagctgtctgcacagctccctgcaagcccctaagtttgggcttgcacagtgtccag	124		
Db	70 CTCCAAGCTGTGTGGCACTCCCTCAAGCCCCCTCAGTTGGGCGCTTCCACAGTGCCAG	129		
QY	125 caagcaagccctggctgtgggttagggagactccctcaagcaagcaagccctgagccgca	184		
Db	130 CAGGCAAGCCCTTGCGTGGGGGTAGGGGACTCCCTTACAGAGCAGCAGCCCTGAGACTCAG	189		
QY	185 ccagcaaccctctgaggtgtgccaagcccccaagvtggccaacctgagtgtgctctgcca	244		
Db	190 AAGGCAACCCCTTGTAGGGGGTGGCCAGGCCCCCAGTGGCCAACTGTGATGCTCTGTGCA	249		
QY	245 ccagccctgctgtgcccctgtgttcgcgtctgccccccagagtccctgtgttagaagccagt	304		
Db	250 CCAGCCCTGTGTGGCCCTGTGTTCCCTTGCGTCCCTCCAGATGCTTGCTGAGACGCCAGT	309		
QY	305 ggactcaagctgccaacacctcttcccgggccccttggagtgtgactgtcagcaagacagctcc	364		
Db	310 GGCCCTCAGGTGCCCACTCTTCCCGGCCCCCTGAAAGTGGCAGTCGACAGACAGCTCC	369		
QY	365 ctgggcaacagagcaactaaacagacagcgccgcaaccccaaaagacagagcgcatgtggcgc	424		
Db	370 CTGGGCACCAAGGCACTAACAACAGACGCGGCACCCAAAGAGAGGCGCATGGGGCAAC	429		
QY	425 gccaaccccggtgtctagcagcglatcccccaacacactctcctgtccccccgacacagtg	484		
Db	430 GCCACGCCGGGTCTGAGCAGCGCTATCCCCACACACTCTCTCTGCCCCCGCACAGGTG	489		
QY	485 tcggcggaacaggtcttggaaagcgcgacggtgggtgcaagtgtgtctcgaaagacaggaactgg	544		
Db	490 TCGCGGACAGGCTTTGGAGMAGCGCGACGCGGGGCAAGTGGGTCTGAGAGAGGGGAACTGG	549		
QY	545 agtcccaagtcacaccgc	604		
Db	550 AATCCCACTCCACCCGCCACCGCCCAAGAGAGGCGCTGTGCGCTTTACTCTCTCTACTTT	609		
QY	605 gacatgtctgtaaccttgaagacagcagcttgggaagccaagcccccttgggtgcagacagtcgg	664		
Db	610 GACATGCTGTACCTTGAGAGCAGCAGACTGGGACGCCAAGGCCCTTGGGGCCAGCAATGCG	669		
QY	665 gaggagggcaacctgacacagctctgaagtgccccggtcatattgaagccaagcccccaaggggc	724		
Db	670 GAGGAGCCACCTGAGAGGCTTGAAGCATGTCCCGGTATTGAAAGCCAAAGCCCCACAGGGGC	729		
QY	725 accctggaactgtgtcccgcggtgtgacacttggagggagacactcgcttggagcaagttggag	784		
Db	730 AACTCGGAATTTGTGTCGGGGGGGTGACCTTGGAGAGCAATCCCTGTGAGAGGATGCGAG	789		
QY	785 tccatgtgtgtgtggcgaaagtgtctcaagagacatcgaaagcgcttcaagctgtctcaatc	844		
Db	790 TCCATGTGTGTGGCGAAATGTCTCAAGAGCAATCGAGACGCGCTTGCAAGTCTCTCAACATC	849		

QY	845	acgcgagatcccatgtgactgagcccccagcaatgtgtcagaaagtgtctcctctgtgacagag	904
Db	850	ACCOCAGATCCCAAGGACATGGAGCCCGACGAAAGTGCACAAATGGCTCCTGTGGACAG	909
QY	905	cacccaataccggtctgtcccccctctggccaagagccttcacagagcttgccggcagaagagt	964
Db	910	CACCAATACCGGCTGCGCCCGCATGGCCAGAGGCTTCACAGAGCTGGCGGCAAGAGACTG	969
QY	965	tcgcgcattgtcggagagagcagttccgcgcagcgtctgcgccttggtgtgggattgtgtctac	1024
Db	970	TGCGCGATGTGGAGAGAGAGTTCCGCGACAGCTCCGCCCTGGGTGGGGATGTCTGTCCAC	1029
QY	1025	gcccaacctgagacatctgaagttaacgagcgtctgattataaagacgaacttaaccttggg	1084
Db	1030	GCCCACTGGACATCTGGAAAGTCAAGCGGCTGGATATAAAGCGACATTCACCTG6GGG	1089
QY	1085	attcaactactgtgtcctclogaccagltgaagagagcttgagaccgaagcagagtgtagctatca	1144
Db	1090	ATTCACTACTGTGGCTCGACACAGAGAGAGAGTGTGACCGACAGGAGGTGACATCA	1149
QY	1145	tgctccggcgagcccataccaacctgtgacagttcctcaagaaattgtctactcaagcccac	1204
Db	1150	TGCTCCGGGCGAGCCCATCCACTGTGGCAATTCCTCAAGAGTCTCACTCAAGCCCAC	1209
QY	1205	agctatgcccctctcatatgaigtgtgtctaaagaagagaagagcatcttcaaatitgaagac	1264
Db	1210	AGCTATGGCCGCTTCATTAGTGGGTCTAACAGAGAAAGGCAATCTTCAAAATTAGGAC	1269
QY	1265	tcaagccaggtgtgcccggctgtgtgggcatccgcagaaccggtccgcgcatagtacgtgac	1324
Db	1270	TCAAGCCAGGTGGGCGGCGGTGGGGGCAATCCGAAACCGTCGCGCCATGAAATCAAC	1329
QY	1325	aagctgaagcgtctcatccgcgcagctcttaagaagaaggacatccggagacacaacatc	1384
Db	1330	AAGCTGAGCCGCTCATTCGCGCAGTATTACAAAGAGCGCATATCCGGAAGCCACATCC	1389
QY	1385	tcccaagcgcctgtctataccagttcgtgtgacaccacatctgagtgcttggcccagggcctgaa	1444
Db	1390	TCCCGAGCCTCTGCTACACAGTTCTGTCAACCCCATGTGATGCTGTGGCCCAAGGCTGAA	1449
QY	1445	accgcgcctagggagcctctctctctgtcgtcctctgtcctagcagaagcccttgatagtg	1504
Db	1450	ACCCGCTCTCAAGGGGCGCTCTCTCTGTGCTCCCTGTGCTCAAGCGGCTCTGAGATGGGG	1509
QY	1505	aaaaagggcagltgtgtctgtcgtcctcagaccttcacagagcccaaggttcagaggaaggagaa	1564
Db	1510	AAAAGGGGCAATGTGCTGTGCTGTGACTTCAGAGGCCAAGGTCAAGGAGGGGCA	1569
QY	1565	ccaactgcccaggggagatatgtgtctctcgtgggaccttgaggacatatgggcagaggtgc	1624
Db	1570	CCAACTGGCCCAAGGGGATAGGGTCTCTGTGGGGCTTGCGGACCTGTGGGCAAGGGGTGC	1629
QY	1625	ttcctcctcagagccagcgtgtctccctctggagagacagaggagagacagggctgtctcccaac	1684
Db	1630	TTCCTCTCTCAGGCCAGCTGCTCCCTCTGAGAGACAGAGGAGAGAGCGCTGCTCCCAAC	1689
QY	1685	acctgtccttgaaccocagcatttcacagaagcagcgtacagaagggcagltgactcgacaa	1744
Db	1690	ACCTGCTCTGACCCCGACATTTCTCAAGACAGAGCTTACAGAAAGGCACTGTGATCGA	1749
QY	1745	aaggcacaaggagatccagagcctctctgtctcatccctcctgtcctccattctgacacac	1804
Db	1750	AGGCAACAGGAGTCCAGGCTCTCTGTGCTCATCCCTGTGCTCCATTCTGACACAC	1809
QY	1805	acctgtgcatgtgtcagggagacatctgcacccctcagtttggcagccagagagltgccccg	1864
Db	1810	ACCTGTGATGTGTCAAGGAGACATCTCAACCTCTGAGTTGGCAGCCAGAGATGCCCCG	1869
QY	1865	ggaatggaataaagaatactagagaactg	1894
Db	1870	GGAAATGAATAAAGATCACTAGAGAACTG	1899

RESULT 6  
AB031549 1898 bp mRNA linear pri 20-JAN-2000  
LOCUS Homo sapiens PSE mRNA for prostate ets, complete cds.  
DEFINITION AB031549  
ACCESSION AB031549.1 GI:6721497  
VERSION prostate ets.  
KEYWORDS prostate ets.  
SOURCE Homo sapiens adult male prostate epithelium prostate cancer  
cell\_line:PC-3 CDNA to mRNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Yamada,N., Tamai,Y., Miyamoto,H. and Nozaki,M.  
TITLE Cloning and expression of the mouse Pse gene encoding a novel Ets  
family member  
JOURNAL Gene 241 (2), 267-274 (2000)  
MEDLINE 2037504  
REFERENCE 2 (bases 1 to 1898)  
AUTHORS Nozaki,M.  
TITLE Direct Submission  
JOURNAL Submitted (25-AUG-1999) Masami Nozaki, Osaka University, Institute  
for Microbial Diseases: 3-1, Yamadaoka, Suita, Osaka 565-0871,  
Japan (E-mail:mnnozaki@iken.osaka-u.ac.jp, Tel:81-6-879-8338,  
Fax:81-6-879-8339)  
FEATURES  
source  
1. .1898  
location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="6p21.2-21.3"  
/sex="male"  
/cell\_line="PC-3"  
/cell\_type="prostate cancer"  
/tissue\_type="prostate epithelium"  
/dev\_stage="adult"  
/note="vector:pUC18"  
406..1413  
/gene="PSE"  
406..1413  
/gene="PSE"  
/codon\_start=1  
/product="prostate ets"  
/protein\_id="BA089543.1"  
/db\_xref="GI:6721498"  
/translation="MGSASPGLSVSPSHLLPDTVSRTGLEKAAAGAVLERDWS  
PSPATPEOGLSAFYSLYDFMDLYPEDSNAAKAPGASSEEPPEDEGPIIDSOAPA  
GSLDIPGGLTEHSLEQVSMVGEVKDIETACKLINTADPMDSPSNVOKML  
KTEHOYRLPMPKARPELAKELCAASEQFRORSPLGSDVLAHLIDTKSAAMKER  
TSPGAIHYCASTSESWTDSVDSSGQPIHMDQPLKELKLPKSHYGRFTIMLNKEK  
GIPIEDSAQAVARLWGIWIRKNRPAMNYDKLSIRQYKGIIRKPIDISORLYOVFVHP  
I"

BASE COUNT 380 a 651 c 567 g 300 t  
ORIGIN  
Query Match 95.7%; Score 1811.8; DB 9; Length 1898;  
Best Local Similarity 98.9%; Pred. No.1.6e-293;  
Matches 1866; Conservative 0; Mismatches 17; Indels 4; Gaps 4;

QY 8 ttctctccagcaattctctgacatctgcgltgccaaacatgcgcccccaagccagtcctc 67  
DB 1 TTCTCTCCAGACATTCCTGACCTGCGTGTGCACACATGCCCCACAGACCAAGTCTCTC 60  
QY 68 caagcctctgcaagcctcctcctgcaagcctcaggttgagccttgcaagtgcaagcaag 127  
DB 61 CAAAGCTGCTGCGACGCTCCTCGCAAGCCCTCAAGTTGGGCTTTGCCACGCTGCGACAG 120  
QY 128 gcaagcctctgagcttgaggttaggagactcctacagcagcagccttgagaccgcaacca 187  
DB 121 GCAAGCCTTGAGGCTGAGGGAGTCTCTACAGGACGACGCTCGAAGACCTCAGAGG 180

QY 188 gccaccccttgaggtgagcgaagcccccaagtgagccaaactgaagtgcctctgcaccaca 247  
DB 181 GCCACCCCTTGAGGGTNG-CAGGCCCCAGTGCCCAACCTGAGTGTCTCTGTCACACA 239  
QY 248 gccctctgccccttgatcgccttgccccccagaatgctctgactgaagacagcagtgagc 307  
DB 240 GCCCTGCT-GCCCCGTGTTCCCTGCTGCCCCCAGATGCTGTGCTGAGACAGCCAGT-GC 297  
QY 308 ctcaagctgccacaactctctccgagcccttgaggttgacatgcacagcagaagcctccctg 367  
DB 298 CTCAGCTGCCACACACTCTTCCCGGCCCTTGAAHTGGACAGTCAGACAGAGCTCCCTG 357  
QY 368 ggcacacagcagctacaagacacagccgcgaagcccccaacagcagcggaatggagcgcc 427  
DB 358 GGCACACGACACTTAACACACACACCCGACCCCAAAACAGCAGCGGATGGGCGCC 417  
QY 428 agcccggtctgagcagcgatcccccaagcaccctcctgctccccccgagacagtgctcg 487  
DB 418 AGCCCGGCTGTAGACAGCTATCCCCAGCCACCTCTCTGCTGCCCCCGACACGGTGTG 477  
QY 488 cgaacagcttgagaaagcgagcgagcgagtgaggtctcgagagaacggaactgagat 547  
DB 478 CGGACAGCTTGAGAGAGCGGCGAGCGGGGCGAGTGGGCTGAGAGAGCGGAGCTGAGT 537  
QY 548 cccaatgcaccccgcaagcccgagcagagggctgtccgctctgcacctctcactttgac 607  
DB 538 CCCAGTCCACCCGACACGCGCCGACGAGGCGCTGTCCGCTTCTACCTCTCTACTTTTAC 597  
QY 608 atgctgtacccttgagacagcagctggtgagcagcagcagcccttggtggcagagatcgagg 667  
DB 598 ATGCTGTACCTTGAGAGACGACAGCTGGGAGCAAGAGCCCTGGGGCCAGAGTGGAGAG 657  
QY 668 gagcacccttgacaagccttgagcagtgcccggtcatatgaacgcaagcccccgagcgacc 727  
DB 658 GAGCCACCTGAGAGAGCGCTGAGCAGTGCCTGGTCATTTGACAGCAGCCCGAGCGGAGC 717  
QY 728 ctgagacttggtcccgagcggtgacttgagagagcactgcgtgagcaggtgactgcc 787  
DB 718 CTGGAATTCGTCCCGCGCGCTGACCTTGAGAGGACCTCCCTGAGACAGTGCAGTCC 777  
QY 788 atggtgtgtggcgaagtgtctcaagacatcgagacgagcgtgcaagctgtctaacaatcac 847  
DB 778 ATGGTGTGTGGGGAAGTGTCTCAAGGACATCGAGAGGCGCTGCAAGCTGTCAACTAC 837  
QY 848 gcaagatcccaatgagcttgagcccccaagcaatgtgcgaagatggtctctgttgacaagagac 907  
DB 838 GCAAGATCCCATGAGATGAGGCCGCCAGCAATGTGCAAGATGTGCTCTGACACAGACAC 897  
QY 908 caataccggtctcccccaatggcagagcctccagagagcgtggcgagcaaggaagctgtgc 967  
DB 898 CAATACCGGCTGCCCCCATGGGCAAGGCGCTTCCAGAGAGCTGGCGGCAAGAGCTGTGC 957  
QY 968 gcaatgtcgaaggaagcagctccgcagcagcgtccgcccctgggttgaggagtgctgcagcgc 1027  
DB 958 GCCATGTGCGAAGAGAGATTCGCGCAGCGCTGCGCTGGGTGGGAGATGTCTGACGCC 1017  
QY 1028 caactgagacatctggaagtcaagcggtcgatgtgaagaagcgagactcaactggggcgatt 1087  
DB 1018 CACCTGAGACATCTGGAAGTCAAGCGCGCTGAGTGAAGAAGCGGACCTTCACTGCGCGATT 1077  
QY 1088 cactactgtgctctgaaacagtgagagagcttgaccggaacgaggttgatctcatatgc 1147  
DB 1078 CACTACTGTGCTCGACCGAGTGAAGAGAGCTGGACGAGAGGAGTGGACTCATCATATGC 1137  
QY 1148 tcgagcgaagcccatcacaactgtgagtgatcctcaagaggttgctactcaagcccaagc 1207  
DB 1138 TCGGAGCAACCATCATCTGTCGTCAGATTCCTCAAGAGATGTCTACTCAAGCCACACAGC 1197  
QY 1208 tatggcgctcatataggtgctcaacaaggaaggaaggaatctcaaaaattgaggaactca 1267  
DB 1198 TATGGCGCTTCAATATAGTGGCTCAACAAGAGAGGAGCATCTTCAAAATTGAGAGATCA 1257  
QY 1268 gcccaagtggtcccggtctgtgggcatccgcaagaacggtcccgccaatgaactgagcaag 1327





```

Db 552 ATGAGTCTCTAGTCCACCCGACCCGACGAGGCGCTGTCTCTTTCACCTCTCTTA 611
Qy 601 cttgaaatgtctgtacccctgagagcaagctgagcgaagccctgggccaag 660
Db 612 CTTTAACA---TGATATCCCGACGATACAGCTGGGTGGTCCAAATGCCCGGAGCCCTGTC 668
Qy 661 tcggagagagcagcactgagcagctgagcagctgagcagctgagcagcagcagc 720
Db 669 CGGGAGAGACCAACCCGGAGAGCCCGAGCACTGTCCCTCTTACAGCCAGGCGCTTGG 728
Qy 721 gggagaccctgagctgtgctgcccggcgagctgaccttggagagagcactgctgagcag 780
Db 729 GAGC-----ACGTGGATGAGACACTCGCTAGAGCAGGT 761
Qy 781 gcagttcaatgtgtgtgagcgaagtgtcgaagagacatcgaagagcagcctgcaagctgtccaa 840
Db 762 GCAATCGATGTTGTGGCGGAGGTCCTGAAGATATTGACAGCGGCTTCAGAGCTTCGAA 821
Qy 841 catcacccagatcccatgagctgagcccgagcaatgtgcaagaagtgtgctgtgac 900
Db 822 CATCACAGACACCTCTGGGGACCTGGAGCCCTGTGATACGTGAGAGATGGCTTTATGAGC 881
Qy 901 agagcaccacaataccggtctgccccccatgtggcaagcgtcttccagagctgtgctgcaagaa 960
Db 882 AGAACACCAGTACCGGCTGCTCCAGCAGCAAGGCTTCACAGAGAGCTGGGCGGTAAGA 941
Qy 961 gctgtgagccatgtgtcggagagagcagcttccgcaagcgtcgcgcctgtgtgtgtgtgtgt 1020
Db 942 GCTGTGCCCCATGTCCGAGGACAGTTCCTGTCACCTGTGCACCTTGTGGTGGGAGTACT 1001
Qy 1021 gcaacccaccctgagacatctgaaagctgagcagcctgtgatgaaagagcagcacttcaactgtg 1080
Db 1002 GCATGCCACCTGGACATCTGGAAGTACAGCGGCTGTGATGAAGAGAGAGACCTGCGCTGG 1061
Qy 1081 ggcgaattcaactgtgtccttcgaccagtgagagagagctgagcgaagcagcagctgtgactc 1140
Db 1062 GACCTTCTACTACTGCGCTCCACAGAGAGAGGCGGAGCGATGGTGAAGTGGAGTGCATC 1121
Qy 1141 atcaagtcggggggagccacatccacactgtgacgttctcgaagagagctgtcaactcgaagcc 1200
Db 1122 GTCTGTGCTCCGGGACGCCCATTCACCTGTGGCACTTCTGAAAGAACTGCTGCTCAAGCC 1181
Qy 1201 ccacagctatgtgcgctcatcattaggtgtgtcacaagaagagagcactcttcaaatgtga 1260
Db 1182 CCACAGCTATGGCGCTTCATTCGCTGCTCAACAGAGAGAAAGCATCTTCAAAATTGA 1241
Qy 1261 ggaactagccaggtgtgcccggctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1320
Db 1242 GGACTCAGCAGAGGTGGCCCGACGTGGGTGTGCGCAAGAACCGCCAGCATGAACCTA 1301
Qy 1321 ggaacagctgagcagctcaccacccagcttacaagaagagcactcaccagagagcaga 1380
Db 1302 TGATTAATTAAGCCCTCATTCGCCCAATATTACAAGAGAGGCAATCAATCTTAACCCGA 1361
Qy 1361 catctccagcgcctcgtcactacagcttgcaccccaactctgagctgac---tggccaag 1436
Db 1362 CATCTCTCAGGCGCTTGTCTACCAATTGTGATCACTGTGAGAGACACAGAGACACAGA 1421
Qy 1437 ggcctgaaacccgcctcgaagggtcctctctctctctctctctctctctctctctctctct 1481
Db 1422 GGCCTACAACTGCCCCAGGACCACTCTGTGGTGGCTGCTGTC 1466

```

```

RESULT 8
AB019436 1704 bp mRNA linear ROD 20-JAN-2000
LOCUS Mus musculus mRNA for Prostate specific transcription factor ets,
DEFINITION complete cds.
ACCESSION AB019436
VERSION AB019436.1 GI:4760536
KEYWORDS Prostate specific transcription factor ets; mpse.
SOURCE Mus musculus (strain:C57Bl/6) Adult male Colon Epithelial CDNA to

```

```

ORGANISM Mus musculus
REFERENCE Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.
AUTHORS Yamada,N., Tamai,Y., Miyamoto,H. and Nozaki,M.
TITLE Cloning and expression of the mouse pse gene encoding a novel Ets
family member
JOURNAL Gene 241 (2), 267-274 (2000)
MEDLINE 20137504
REFERENCE 2 (bases 1 to 1704)
AUTHORS Nozaki,M.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-1998) Masami Nozaki, Osaka University, Institute
for Microbial Diseases, 3-1, Yamadaoka, Suita, Osaka 565-0871,
Japan (E-mail:mozaki@ibiken.osaka-u.ac.jp, Tel:81-6-879-8338,
Fax:81-6-879-8333)
FEATURES
source
1..1704
/organism="Mus musculus"
/strain="C57Bl/6"
/db_xref="taxon:10090"
/sex="male"
/cell_type="Epithelial"
/tissue_type="Colon"
/dev_stage="Adult"
362..1339
/gene="pse"
362..1339
/gene="mpse"
/codon_start=1
/product="Prostate specific transcription factor ets"
/protein_id="BAA77329.1"
/db_xref="GI:4760537"
/translation="MGSAIPGLSNVSPCLILFPVAPRTGTERAASGAMPEKQEW
ESPATPEGLISAATYLSRNMTPDSSWYAKYPPARAGEDHPPEPDQPVYDSAGS
TLDESLRQVOSMVVGVLEKIDETACKLINTADGDSPNVOKWILMTBOYRLPP
AKARQELGKEKLCAMSBEQFNORAPFLGDDVLAHLDTMKSAAMKERTSPGTHYCA
STSEGWTDGEVDSGSCOPILHMQFLKEILLKPSHYGRFIRMLKEKGIKIDBSAQ
VARLVGNVRNRPAMYDKLSIRSIIRQYKKGIIRKPDISORLYVGVHVV"
1704
polyA_site
/note="10 a nucleotides"
BASE COUNT 369 a 527 c 491 g 317 t
ORIGIN
Query Match 39.0%; Score 738.4; DB 10; Length 1704;
Best Local Similarity 73.3%; Pred. No. 4,8e-114;
Matches 1061; Conservative 0; Mismatches 336; Indels 51; Gaps 7;
Qy 38 tgtcacacatgcccacagaccagctcctcacaagcctgtgtccagctccttgcagagcccc 97
Db 1 TGTCCGCTGTGCTCCACACCTAGACCCACCC- -GCTGCTGCCCGGAGAGAACCC 58
Qy 98 tcaagttggccttgcacaggtgtccagcagcagcctgtggtgagtgagggactccc 157
Db 59 CCAGCTTGGGCTTGTCTATGCTGCCAGAGTGGCCCTGAGCTTCAAGAGGCGCTGTC 118
Qy 158 tacagcagcagcagccttgaagcagcagcagcagcagcagcagcagcagcagcagc 217
Db 119 TATAGACTGTGACAGGCTTGAAGCC- -TCAGACTCACTCAAGAGGCGAAGAGCCCTTG 174
Qy 218 tggcacaactgagtgctgtccttgcacacagccttgcagcagccttgcagcagcccc 277
Db 175 TGGCCCACTTAAGAGGACACTGTGTCCCAAGCCCTCTGCGCCACTGAT- -GTCAG 228
Qy 278 ccagatgctgtgtgagacagcagcagcagcagcagcagcagcagcagcagcagcagc 337
Db 229 ACTGAGACCCAGCAGTGAACCC- -TGAGCTGCTGCGCCACTGCTCTCTGTGGTCCCT 283
Qy 338 ggaagttggacactgacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 397
Db 284 GAGGTGTGCTCTGCGAGGACGACGACTTCTTGAAGCAGGCGGCTTAACGGAAGCAGCC 343

```





QY	1304	gctccgcacatgaactctgcacaaagctctggagccctccatccgcgcagctctcaagaagaagc	1363
Dd	123679	cgtcccccctatgaactacgacacagctgagccctccatccgcgaattattacagaagaagcc	123520
QY	1364	atcaatccggaagccagaacatctccagcgcctcgtctcaacagttcgtgcaccccatctga	1423
Dd	123619	atcattcccggaacccagacatctctccagccctcgtctacgaattgtgtgaccccatctga	123560
QY	1424	gtgccttgagccagagccctcaaacccgcctcaaggagccctctcctgtcgtccctgcctc	1483
Dd	123555	gtgccttgagccagagccctcaaacccgcctcaaggagccctctcctgtcgtccctgcctc	123500
QY	1484	agccagcccttgagatctgggggaaaaacgggcagttgtctctgtcgtcctgaacctccagag	1543
Dd	123499	agccagccctctagatctgggggaaaaacgggcagttgtctctgtcgtcctgaacctccagag	123440
QY	1544	cccaaggtcagaggaagggaacaaactctcccaaggaggaataatggctctctctggagccctc	1603
Dd	123439	cccaaggtcagaggaagggaacaaactctcccaaggaggaataatggctctctctggagccctc	123380
QY	1604	ggagccatctggggcaggggtgtctctcctccagcccaagctgtctccctctggagaagaag	1663
Dd	123379	ggagccatctggggcaggggtgtctctcctccagcccaagctgtctccctctggagaagaag	123320
QY	1664	ggagcagaagcctctccccaacacctgcctctcaaccacagcatctccagagcagaagcctac	1723
Dd	123319	gagacagagcctctccccaacacctgcctctcaaccacagcatctccagagcagaagcctac	123260
QY	1724	agaaaggcagtgactctgacaaaagccaaagcagttccagcctctctctctgtcctcatcccc	1783
Dd	123259	agaaaggcagtgactctgacaaaagccaaagcagttccagcctctctctctgtcctcatcccc	123200
QY	1784	ctgcctctccatctgcacacacccctggcagtggtgaggaagacatctgaacccctcagtt	1843
Dd	123199	ctgcctctccatctgcacacacccctggcagtggtgaggaagacatctgaacccctcagtt	123140
QY	1844	ggagcagccaggaagctcccccgggaatgatataataagatactagagaactg	1894
Dd	123139	ggagcagccaggaagctcccccgggaatgatataataagatactagagaactg	123089

LOCUS	AX048155	1751 bp	DNA	linear	PAT 15-DEC-2000
DEFINITION	Sequence 8 from Patent WO0070092.				
ACCESSION	AX048155				
VERSION	AX048155.1	GI:11876971			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1 (bases 1 to 1751)				
AUTHORS	Kaufmann, J. and Xin, H.				
TITLE	Expression of ets-domain proteins in cancer				
JOURNAL	Patent: WO 0070092-A 8 23-NOV-2000;				
	Chilton Corporation (US)				
FEATURES	Location/Qualifiers				
source	1..1751				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
promoter	1..1381				
	/note="JTKETS promoter"				
TATA_signal	1356..1360				
	/note="TATA box."				
BASE COUNT	369 a 521 c 497 g 364 t				
ORIGIN					
Query Match	17.4%	Score 329.8;	DB 6;	Length 1751;	
Best Local Similarity	97.7%	Fid. No. 9.8e-46;			
Matches 345; Conservative	0;	Mismatches 7;	Indels 1;	Gaps 1;	

Accession	Sequence	Length
Oy 1	gttgattctctccagacatctctgtgacatctgcgtgtccaaatcgcccaagacc	60
Db 1400	gttcgactttctccacgacacatctctgcacgtctgccgtgtccacacgtgcccaaaccc	1459
Oy	61 agtctccaaagctgtcgcagatccctctgcaagccctcagttgttgacctgccaagtg	120
Db 1460	agttcttccaaagctgtgcacactctccctgcacaaagccctcagattgtggcccttgccacggtg	1519
Oy	121 ccagcagcagcgcctgtgctgtggttgaggactccctcacaagcagcagccttgagac	180
Db 1520	ccagcagcagcgccttggtgtgggttgggggttgggggacttccctacagcagcagcagccttgagacc	1579
Oy	181 gccacacagccaccttgagggttgccagagccccaattgccaacctgagttgctgcctc	240
Db 1580	tcacagggccaccccttggagggtggccacagcccccacgt-gccaaactcgaattgctgtgcttt	1638
Oy	241 gccacacagcgcctgtgcccctgtgcttcgccttgagcccccagatgcctgtgctgacaagc	300
Db 1639	gccacacagcgcctgtgtggccctctgtgttccgctggcccccaaatgctgtggcttgagacagc	1698
Oy	301 cagtgagcctcagatgcgcccaacctcttcgccggccctcgagattgtgcaatcgag	353
Db 1699	cagtgagcctcagatgcgcccaacctcttcccggccctctgaattggcactctcag	1751

[illegible]

QY 1221 ttagtgctcaacaagaagaagg 1244  
|||||  
Db 255 TTANGTGCTCAACAGAGAGAGG 278

RESULT 12  
AX140908 AX140908 278 bp DNA linear PAT 31-MAY-2001  
DEFINITION Sequence 398 from Patent WO0134802.  
ACCESSION AX140908  
VERSION AX140908.1 GI:14281005  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 278)  
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,  
Skelky,Y.A. and Wang,A.  
TITLE Compositions and methods for the therapy and diagnosis of prostate  
cancer  
JOURNAL Patent: WO 0134802-A 398 17-MAY-2001;  
CORIXA CORPORATION (US)  
FEATURES  
Location/Qualifiers  
1. 278  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 56 a 85 c 87 g 49 t 1 others  
ORIGIN

Query Match 13.9%; Score 263; DB 6; Length 278;  
Best Local Similarity 99.6%; Pred. No. 2.2e-34;  
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 981 agcaattccgacgagctcgccctgggtggatgtgtcgtcgaagccacattgacatct 1040  
|||||  
Db 15 AGCAATTCCGACGAGCGCTCGCCCTGGGTGGGATGTGTCTGACGCGCCCTGACATCT 74

QY 1041 ggaatcagcgagctgtgataagaagcggaattacactggggcgatcactactgtgct 1100  
|||||  
Db 75 GGAATCAGCGGCGCTGTGATGAAGAGCGGACTCAGCTGGGGCGATTCACACTGTGCTCT 134

QY 1101 cgaccagtgaagagagctgtgagccagcagcgagtgactatcactgtctccggcgagccca 1160  
|||||  
Db 135 CGACCAGTGAAGAGAGCTGTGACCGACGAGGTGACTATCTATCTCCGGCGAGCCCA 194

QY 1161 tccacctgtgagcttccctcaagaagtgatctactcaagcccaagctatgagcccttca 1220  
|||||  
Db 195 TCCACCTGTGCGAGTTCCTCAAGAGGTGTGACTCAAGCCCCACAGCTATGCGCGCTTCA 254

QY 1221 ttagtgctcaacaagaagaagg 1244  
|||||  
Db 255 TTANGTGCTCAACAGAGAGAGG 278

RESULT 13  
AX200768 AX200768 278 bp DNA linear PAT 29-AUG-2001  
DEFINITION Sequence 398 from Patent WO0151633.  
ACCESSION AX200768  
VERSION AX200768.1 GI:15390661  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 278)  
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,  
Stolk,J.A., Skelky,Y.A., Wang,A. and Meagher,M.J.  
TITLE Compositions and methods for the therapy and diagnosis of prostate

JOURNAL Patent: WO 0151633-A 398 19-JUL-2001;  
CORIXA CORPORATION (US)  
FEATURES  
Location/Qualifiers  
1. 278  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 56 a 85 c 87 g 49 t 1 others  
ORIGIN

Query Match 13.9%; Score 263; DB 6; Length 278;  
Best Local Similarity 99.6%; Pred. No. 2.2e-34;  
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 981 agcaattccgacgagctcgccctgggtggatgtgtcgtcgaagccacattgacatct 1040  
|||||  
Db 15 AGCAATTCCGACGAGCGCTCGCCCTGGGTGGGATGTGTCTGACGCGCCCTGACATCT 74

QY 1041 ggaatcagcgagctgtgataagaagcggaattacactggggcgatcactactgtgct 1100  
|||||  
Db 75 GGAATCAGCGGCGCTGTGATGAAGAGCGGACTCAGCTGGGGCGATTCACACTGTGCTCT 134

QY 1101 cgaccagtgaagagagctgtgagccagcagcgagtgactatcactgtctccggcgagccca 1160  
|||||  
Db 135 CGACCAGTGAAGAGAGCTGTGACCGACGAGGTGACTATCTATCTCCGGCGAGCCCA 194

QY 1161 tccacctgtgagcttccctcaagaagtgatctactcaagcccaagctatgagcccttca 1220  
|||||  
Db 195 TCCACCTGTGCGAGTTCCTCAAGAGGTGTGACTCAAGCCCCACAGCTATGCGCGCTTCA 254

QY 1221 ttagtgctcaacaagaagaagg 1244  
|||||  
Db 255 TTANGTGCTCAACAGAGAGAGG 278

RESULT 14  
AX267424 AX267424 278 bp DNA linear PAT 26-OCT-2001  
DEFINITION Sequence 398 from Patent WO0173032.  
ACCESSION AX267424  
VERSION AX267424.1 GI:16516188  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (siles)  
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,  
Vedick,T.S., Carter,D., Li,S.X., Wang,A., Skelky,Y.A., Hepler,W.T.  
and Henderson,R.A.  
TITLE Compositions and methods for the therapy and diagnosis of prostate  
cancer  
JOURNAL Patent: WO 0173032-A 398 04-OCT-2001;  
CORIXA CORPORATION (US)  
FEATURES  
Location/Qualifiers  
1. 278  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 56 a 85 c 87 g 49 t 1 others  
ORIGIN

Query Match 13.9%; Score 263; DB 6; Length 278;  
Best Local Similarity 99.6%; Pred. No. 2.2e-34;  
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 981 agcaattccgacgagctcgccctgggtggatgtgtcgtcgaagccacattgacatct 1040  
|||||  
Db 15 AGCAATTCCGACGAGCGCTCGCCCTGGGTGGGATGTGTCTGACGCGCCCTGACATCT 74

QY 1041 ggaatcagcgagctgtgataagaagcggaattacactggggcgatcactactgtgct 1100  
|||||









PF 02-AUG-1999: 99WO-US17470.  
XX  
PR 31-JUL-1998: 98US-0126945.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
PI Liberman TA, Oetting JP, Kunsch CA, Endress GA, Rosen CA;  
XX  
XX  
DR MPI: 2000-195255/17.  
DR P-PSDB: AAY45003.  
XX  
XX  
PT Novel prostate derived polypeptide, polynucleotide useful for  
PT diagnosis, prevention and treatment of prostate cancer, autoimmune  
PT disorders, microbial infections and also as food additive or  
PT preservative -  
XX  
PS Claim 4; Fig 1; 132pp; English.  
XX  
XX The present cDNA sequence encodes a human prostate derived Ets factor  
CC (PDEF), isolated from human prostate carcinoma cDNA library  
CC (ATCC No.203072). PDEF gene expression is highest in tissues involved in  
CC androgen sensitivity, with enriched expression in prostate epithelium  
CC cells. PDEF gene is mapped to the human chromosome 6p21.3  
CC region that is associated with loss of heterozygosity and chromosomal  
CC translocations in various human cancers. PDEF has cytoskeletal,  
CC immunosuppressive, cerebroprotective, fungicide, antibacterial,  
CC antiviral, neuroprotective, antiparkinsonian, nootropic, anabolic,  
CC antiinflammatory and anorectic activity. PDEF polynucleotides are useful  
CC in linkage analysis as markers, as hybridisation probes for differential  
CC identification of the tissues or cell types and as polymorphic markers  
CC for forensic purposes. PDEF is useful as prostate-specific tumour marker  
CC for the diagnosis and treatment of prostate cancer. PDEF sequences are  
CC useful for treating autoimmune disorders, haematopoietic, blood  
CC coagulation, immune and nervous system disorders, hyperproliferative  
CC disorders like, neoplasms and microbial infections, heart attacks,  
CC stroke, scarring and for tissue regeneration. They are also useful as  
CC food additives or preservatives.  
XX  
XX  
SQ Sequence 1894 BP; 368 A; 653 C; 571 G; 302 T; 0 other;

Query Match 98.7%; Score 1870; DB 21; Length 1894;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1879; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 gctgactctctccagacatctctgcaactctgctgctgctccacactgcccagagacc 60  
Db 1 gctgactctctccagacatctctgcaactctgctgctgctccacactgcccagagacc 60  
QY 61 agtctccaaagctgtgcaagctccctcgaagccctcgaagtgtggctgtgccaagtg 120  
Db 61 agtctccaaagctgtgcaagctccctcgaagccctcgaagtgtggctgtgccaagtg 120  
QY 121 ccagcagagcagcctgtgctggtggtgtagggatctccctacagcagcagccttgagcc 180  
Db 121 ccagcagagcagcctgtgctggtggtgtagggatctccctacagcagcagccttgagcc 180  
QY 181 gcaaccagcacccttgagggtggtgcaagcccccaagtgccaacttgatgtgctctc 240  
Db 181 tcagtagggcaccaccttgagggtggtgcaagcccccaagtgccaacttgatgtgctctc 240  
QY 241 gcaaccagcctgtgccccctgttcgcgtgccccccagatgctgtgctgagaaagc 300  
Db 241 gcaaccagcctgtgccccctgttcgcgtgccccccagatgctgtgctgagaaagc 300  
QY 301 cagtgagctagctgcccacaactcttccggccttgtagtggcactgcaagcagaag 360  
Db 301 cagtgagctagctgcccacaactcttccggccttgtagtggcactgcaagcagaag 360  
QY 361 ctccctgggacagagcaggtaaacagacacagcgcacaccccaaaagagcgcatgg 420  
Db 361 ctccctgggacagagcaggtaaacagacacagcgcacaccccaaaagagcgcatgg 420

QY 421 cagcgccagcccggtctgagcagcgtatcccccagccacctctgctgccccccagacac 480  
Db 421 cagcgccagcccggtctgagcagcgtatcccccagccacctctgctgccccccagacac 480  
QY 481 ggtgtcgagacagccttgagaaagcgagcggggtgagctggtctcgaagacggga 540  
Db 481 ggtgtcgagacagccttgagaaagcgagcggggtgagctggtctcgaagacggga 540  
QY 541 ctggaagctccagctccaccgcgcagcccgagcagggcctgtgcgccttcgagccctcta 600  
Db 541 ctggaagctccagctccaccgcgcagcccgagcagggcctgtgcgccttcgagccctcta 600  
QY 601 ctctacatgctgtaaccttgagagacagcagctgggacagcccaagcccttggtgccaag 660  
Db 601 ctctacatgctgtaaccttgagagacagcagctgggacagcccaagcccttggtgccaag 660  
QY 661 tcgggaagagccacactgagcagccttgagcagctgcccgtatctgaacgaagcccaagc 720  
Db 661 tcgggaagagccacactgagcagccttgagcagctgcccgtatctgaacgaagcccaagc 720  
QY 721 gggcagccctggaactgtggtcccgcggtgagccttgagagagacactgcgtgagcaggt 780  
Db 721 gggcagccctggaactgtggtcccgcggtgagccttgagagagacactgcgtgagcaggt 780  
QY 781 ggaagtcagtgtgtggtggtgagagctgctcaagagacatcgaagcggccttgcaagctgca 840  
Db 781 ggaagtcagtgtgtggtggtgagagctgctcaagagacatcgaagcggccttgcaagctgca 840  
QY 841 catcacgcagatcccaatgagctgagagcccaagcaatgtcagaagtgtcctctgtgac 900  
Db 841 catcacgcagatcccaatgagctgagagcccaagcaatgtcagaagtgtcctctgtgac 900  
QY 901 aagagccaataccggctgcgcccccatggtgcaagccttcaggagctgggtggccaagga 960  
Db 901 aagagccaataccggctgcgcccccatggtgcaagccttcaggagctgggtggccaagga 960  
QY 961 gctgtgcgcaatgtcgagagcagcttcgcgcagcgtcgccttggtgtgggtgtgtgt 1020  
Db 961 gctgtgcgcaatgtcgagagcagcttcgcgcagcgtcgccttggtgtgggtgtgtgt 1020  
QY 1021 gcaagcccaactgagacatctggaagtcaagcggctgtggtgaagaagcggatctaactgtg 1080  
Db 1021 gcaagcccaactgagacatctggaagtcaagcggctgtggtgaagaagcggatctaactgtg 1080  
QY 1081 ggcgaattactactgtgctcgcagcagctgagagagcgtgagccagcagcgaggtgagctc 1140  
Db 1081 ggcgaattactactgtgctcgcagcagctgagagagcgtgagccagcagcgaggtgagctc 1140  
QY 1141 atcatgtctccggagcagccatccacactgtggaagcttcctcaagagatgtctactcaagcc 1200  
Db 1141 atcatgtctccggagcagccatccacactgtggaagcttcctcaagagatgtctactcaagcc 1200  
QY 1201 ccaagcgtatggtgcgcttcattaggtgtgtctcaaaagaagaagggatcttcaaatgtga 1260  
Db 1201 ccaagcgtatggtgcgcttcattaggtgtgtctcaaaagaagaagggatcttcaaatgtga 1260  
QY 1261 ggaactcaagccaggtggtgcggctgtgtgggcatccgcaagaagcgttcccgctcaactgtg 1320  
Db 1261 ggaactcaagccaggtggtgcggctgtgtgggcatccgcaagaagcgttcccgctcaactgtg 1320  
QY 1321 cgaacaagctgagcgcgtccatcagcgcagtatataagaagaaggcatcattccgaaagcaga 1380  
Db 1321 cgaacaagctgagcgcgtccatcagcgcagtatataagaagaaggcatcattccgaaagcaga 1380  
QY 1381 catctccagagcgtcgtgtactacagatctgtgacaccccatctgaatgtcgtgccaagggcc 1440  
Db 1381 catctccagagcgtcgtgtactacagatctgtgacaccccatctgaatgtcgtgccaagggcc 1440  
QY 1441 tgaacccggcctcaagggtctctctgtgctgacctgagccagcggcctgtgagtgtg 1500  
Db 1441 tgaacccggcctcaagggtctctctgtgctgacctgagccagcggcctgtgagtgtg 1500



```
OY 1501 ggggaaacgggagtgctgtctgtctgtacacctccagaagcccaaggtcaaggaggg 1560
DB 1501 ggggaaacgggagtgctgtctgtctgtacacctccagaagcccaaggtcaaggaggg 1560
OY 1561 gcaacacatgccccaggggagatggtctctgtgggcttcggagccatgggcaaggg 1620
DB 1561 gcaacacatgccccaggggagatggtctctgtgggcttcggagccatgggcaaggg 1620
OY 1621 ggtcttctctctcagggccagctgtctcccttggaagacagagagagagagctctcc 1680
DB 1621 ggtcttctctctcagggccagctgtctcccttggaagacagagagagagagctctcc 1680
OY 1681 caaacacctctctgtgccccagcatcttcagagacagagccctcagaaggagtgactgg 1740
DB 1681 caaacacctctctgtgccccagcatcttcagagacagagccctcagaaggagtgactgg 1740
OY 1741 acaaaaggcacaagcgagctcagagcctctctgtctccatcccccctgcctccatctgca 1800
DB 1741 acaaaaggcacaagcgagctcagagcctctctgtctccatcccccctgcctccatctgca 1800
OY 1801 ccacacctgcatgtgtcagagagacatctgcaacccctagttgggagccagagagtgcc 1860
DB 1801 ccacacctgcatgtgtcagagagacatctgcaacccctagttgggagccagagagtgcc 1860
OY 1861 cccgggaatggaataaagatactagagaactg 1894
DB 1861 cccgggaatggaataaagatactagagaactg 1894

RESULT 2
AAH47005
ID AAH47005 standard; cDNA; 1894 BP.
XX
XX AAH47005;
XX
XX 29-OCT-2001 (first entry)
XX
XX Human ETS suppressor factor (esf) protein encoding cDNA.
XX
XX ETS suppressor factor; esf; apoptosis; cancer; cytostatic; human;
XX growth-suppressor; transcription factor; gene-therapy; ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 416..1423
XX FT /tag= a
XX FT /product= "human esf"
XX
XX WO200157186-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US03640.
XX
XX 04-FEB-2000; 2000US-0180244.
XX
XX (MUSC-) MUSC FOUND RES DEV.
XX
XX Watson DK;
XX
XX WPI: 2001-465703/50.
XX P-PSDB; AAB83609.
XX
XX
XX Fragment of a human ETS suppressor factor protein for inhibiting the
XX growth or proliferation of and/or inducing apoptosis of cancer cells,
XX especially prostate, breast, colon or cervical cancer cells for
XX treatment.
XX
XX Examples: Fig 1A; 94pp; English.
XX
XX The invention relates to methods for treating or preventing cancer by
XX modifying the expression of ETS suppressor factor (esf) protein gene or
```

```
CC the activity of the gene product. The esf protein or its modified form
CC are useful for inhibiting the growth or proliferation of and/or inducing
CC apoptosis of cancer cells, especially prostate, breast, colon or cervical
CC cancer cells. They are useful for treating cancer in a subject when
CC introduced in to the vicinity of the cancer in a subject. The methods
CC provided are useful for inhibiting the specific degradation of esf
CC protein in a cell; for diagnosing a preneoplastic or neoplastic condition
CC in a human by e.g. PCR and for identifying a compound that inhibits the
CC degradation of esf protein in cancer cells. The present sequence
CC represents a human esf protein encoding cDNA.
XX
XX Sequence 1894 BP; 368 A; 653 C; 571 G; 302 T; 0 other;
XX
XX Query Match 98.7%; Score 1870; DB 22; Length 1894;
XX Best Local Similarity 99.2%; Pred. No. 0;
XX Matches 1879; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
OY 1 gctgactctctccagacatctctgtacactctgcgtgtcaacatgccccaaagacc 60
DB 1 gctgactctctccagacatctctgtacactctgcgtgtcaacatgccccaaagacc 60
OY 61 agtctccaagctgtgctcagctccctcctgcaagccctcaggttgggcttgccaaggtg 120
DB 61 agtctccaagctgtgctcagctccctcctgcaagccctcaggttgggcttgccaaggtg 120
OY 121 ccagcagcagcagcctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 180
DB 121 ccagcagcagcagcctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 180
OY 181 gcaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 240
DB 181 gcaacacacacacacacacacacacacacacacacacacacacacacacacacacacac 240
OY 241 gccaccagcctgtgctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 300
DB 241 gccaccagcctgtgctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 300
OY 301 cagtgtcctcagctgtgctcagcctctctccgagcccttgagattgtgcaactgagacagac 360
DB 301 cagtgtcctcagctgtgctcagcctctctccgagcccttgagattgtgcaactgagacagac 360
OY 361 ctccctgtgacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420
DB 361 ctccctgtgacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420
OY 421 cagcgcagcgcgcgcgtgtgagcagcgtatcccccagcaccctctgtggtggtggtggtggt 480
DB 421 cagcgcagcgcgcgcgtgtgagcagcgtatcccccagcaccctctgtggtggtggtggtggt 480
OY 481 ggtgtcgcgcgcgcgtgtgagcagcgtgagcagcgtgagcagcgtgagcagcgtgagcagc 540
DB 481 ggtgtcgcgcgcgcgtgtgagcagcgtgagcagcgtgagcagcgtgagcagcgtgagcagc 540
OY 541 ctgagatccagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 600
DB 541 ctgagatccagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 600
OY 601 ctttgcacatgctgtacccctgagacagcagctgtgagcagcagcagcagcagcagcagcag 660
DB 601 ctttgcacatgctgtacccctgagacagcagctgtgagcagcagcagcagcagcagcagcag 660
OY 661 tcggagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
DB 661 tcggagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
OY 721 gggcagcctgactgtgtgcccggggtgacttggagagacatctgctgagagcaggt 780
DB 721 gggcagcctgactgtgtgcccggggtgacttggagagacatctgctgagagcaggt 780
OY 781 gcaatcagctgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 840
DB 781 gcaatcagctgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 840
```

QY 841 catcacgcagatcccaatggaatcgaagcccaagcaatgtgcaagaatggtctgtgac 900  
 Db 841 catcacgcagatcccaatggaatcgaagcccaagcaatgtgcaagaatggtctgtgac 900  
 QY 901 agagacacacaaataacggtgtgcccccaatggcgaagagcttccagagatgtgcggcaagga 960  
 Db 901 agagacacacaaataacggtgtgcccccaatggcgaagagcttccagagatgtgcggcaagga 960  
 QY 961 gctgtgcacatgctcggagagagagatctccgcacagctgcgccccctgtgggtggagatgct 1020  
 Db 961 gctgtgcacatgctcggagagagagatctccgcacagctgcgccccctgtgggtggagatgct 1020  
 QY 1021 gcaacgcacacatggaatcgtgaagtcagcgcgtgtggaatgaagaagagagagatcactcgtg 1080  
 Db 1021 gcaacgcacacatggaatcgtgaagtcagcgcgtgtggaatgaagaagagagagatcactcgtg 1080  
 QY 1081 ggcgaatcactactgtgtctctgacacagtgagagagatgtgacacgcagcagagtgagatc 1140  
 Db 1081 ggcgaatcactactgtgtctctgacacagtgagagagatgtgacacgcagcagagtgagatc 1140  
 QY 1141 atcatgtctcgcgagcagccacatccacatctgagcaatctcctcaagagatgtgactcaagcc 1200  
 Db 1141 atcatgtctcgcgagcagccacatccacatctgagcaatctcctcaagagatgtgactcaagcc 1200  
 QY 1201 ccacagctatgagcagcttcaatgaagtgtgctcaacaagaagagagagagatcttcaaaattga 1260  
 Db 1201 ccacagctatgagcagcttcaatgaagtgtgctcaacaagaagagagagagatcttcaaaattga 1260  
 QY 1261 ggaactcagccagagtgagcgcgctgtgtgggagatctcgcgaagaagacgctccgcgaatga 1320  
 Db 1261 ggaactcagccagagtgagcgcgctgtgtgggagatctcgcgaagaagacgctccgcgaatga 1320  
 QY 1321 cgacaagctgagcgcgcacatccacatccagcttcaacaagaagagagatctcgcgaagacaga 1380  
 Db 1321 cgacaagctgagcgcgcacatccacatccagcttcaacaagaagagagatctcgcgaagacaga 1380  
 QY 1381 catctccagcgcctctgcttaacagttcgtgcaaccacatctgagctgtgcggcccaagggcc 1440  
 Db 1381 catctccagcgcctctgcttaacagttcgtgcaaccacatctgagctgtgcggcccaagggcc 1440  
 QY 1441 tgaaccgcgcgcctcaggggctctctctctgctgtgcctgtcgaagcagcccttgagatg 1500  
 Db 1441 tgaaccgcgcgcctcaggggctctctctctgctgtgcctgtcgaagcagcccttgagatg 1500  
 QY 1501 ggggaaaaacgggcagtgctgtctgtctgtctgtcgtcgtcgaagcagcccttgagatg 1560  
 Db 1501 ggggaaaaacgggcagtgctgtctgtctgtcgtcgtcgtcgaagcagcccttgagatg 1560  
 QY 1561 gcaacacacatgccccagggagataltggtctctgtgagccttcgagccatggggcaggg 1620  
 Db 1561 gcaacacacatgccccagggagataltggtctctgtgagccttcgagccatggggcaggg 1620  
 QY 1621 gtgtctctctctcagcgcgcgtgtct 1680  
 Db 1621 gtgtctctctctcagcgcgcgtgtct 1680  
 QY 1681 caaacacctgctctgacccacacatctcagagcagagagcctacagaagggcagtgactcg 1740  
 Db 1681 caaacacctgctctgacccacacatctcagagcagagagcctacagaagggcagtgactcg 1740  
 QY 1741 acaaaaggcacaagagcagtcacagcctctctctctctctctctctctctctctctctctctct 1800  
 Db 1741 acaaaaggcacaagagcagtcacagcctctctctctctctctctctctctctctctctctctct 1800  
 QY 1801 ccaaacctgacatgtgtcagagagagacatctgcacccctcagtggtggagcagcaagatgccc 1860  
 Db 1801 ccaaacctgacatgtgtcagagagagacatctgcacccctcagtggtggagcagcaagatgccc 1860  
 QY 1861 ccccggaatgataataaagataactaagaactg 1894  
 Db 1861 ccccggaatgataataaagataactaagaactg 1894

RESULT 3  
 AAC83261  
 ID AAC83261 standard: DNA; 1894 BP.  
 AC AAC83261;  
 XX  
 DT 16-MAR-2001 (first entry)  
 DE Gland-specific Ets transcription factor (GSEF) nucleotide sequence.  
 XX  
 KW Transcription factor; gland-specific Ets transcription factor; GSEF;  
 KW metastatic potential; cancer; tumour; metastasis; breast; prostate;  
 KW leukemia; lymphoma; sarcoma; melanoma; chromosome 6p21.1-6p21.3; ds.  
 OS Homo sapiens.  
 XX  
 PN WO200070092-A1.  
 PD 23-NOV-2000.  
 PF 12-MAY-2000; 2000MO-US13173.  
 PR 14-MAY-1999; 99US-0134112.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Kaufmann J, Xin H, Harrowe G;  
 DR WPI: 2001-041019/05.  
 DR F-PSDB; AAB49628.  
 XX  
 PT Detecting metastatic and potential metastatic cancerous cells, useful  
 PT for diagnosing, prognosing, grading and staging of cancers by detecting  
 PT gland-specific Ets transcription factor gene product in a biological  
 PT sample from a cell  
 XX  
 PS Disclosure: Page 81-83; 95pp; English.  
 XX  
 CC This invention relates to a method for the detection and determination of  
 CC the metastatic potential of a cell. The method comprises detecting a  
 CC gland-specific Ets transcription factor (GSEF) gene product in a test  
 CC sample. Detection of a GSEF gene product in the test sample in amount  
 CC lower than that in a normal cell, is indicative of a cell with high  
 CC metastatic potential. The method is useful for determining the metastatic  
 CC potential of a cell, for the diagnosis and prognosis of cancer as well as  
 CC grading and staging of cancers by detecting GSEF expression in a  
 CC biological test sample. The method may also be used to monitor patients  
 CC having a predisposition to develop a particular cancer. GSEF polypeptides  
 CC are useful for producing antibodies, in cancer diagnosis, prognosis,  
 CC grading, staging and management of breast and prostate tumours, and in  
 CC detecting polymorphisms in the sequence. GSEF genes and proteins are also  
 CC useful in gene therapy. GSEF gene product expression levels can be used  
 CC in conjunction with any tissue in which an alteration in GSEF gene  
 CC product expression levels is associated with development of a  
 CC cancer-associated phenotype. Cancers, which can be monitored include  
 CC cancers of the prostate, cervix, lung and colon, melanomas, colorectal  
 CC adenocarcinomas, Wilms' tumour, retinoblastoma, sarcomas, myosarcomas,  
 CC lung carcinomas, leukemia, and lymphomas. The GSEF gene is located on  
 CC human chromosome 6, specifically at 6p21.1-6p21.3. The present sequence  
 CC represents the DNA encoding GSEF.  
 XX  
 SQ Sequence 1894 BP; 368 A; 653 C; 571 G; 302 T; 0 other;  
 Query Match 98.7%; Score 1870; DB 22; Length 1894;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 1879; Conservativity 0; Mismatches 15; Indels 0; Gaps 0;  
 OY 1 gctgaactccctccagcacatctcctgtcacctgtgcgtgtccacactgtccacagacc 60  
 Db 1 gctgaactccctccagcacatctcctgtcacctgtgcgtgtccacactgtccacagacc 60

QY 61 agtctccaaagcctgctgcaagctccctctgcaagccctctcaggttgaggcctcttgcaagctg 120  
 |||||  
 Db 61 agtctccaaagcctgctgcaagctccctctgcaagccctctcaggttgaggcctcttgcaagctg 120  
 QY 121 ccagcaagcagccctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 180  
 |||||  
 Db 121 ccagcaagcagccctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 180  
 QY 181 gcaacacagcagccctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 240  
 |||||  
 Db 181 gcaacacagcagccctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 240  
 QY 241 gcaacacagcagccctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 300  
 |||||  
 Db 241 gcaacacagcagccctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 300  
 QY 301 cagtgagcctcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 360  
 |||||  
 Db 301 cagtgagcctcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 360  
 QY 361 ctccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
 |||||  
 Db 361 ctccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
 QY 421 cagc 480  
 |||||  
 Db 421 cagc 480  
 QY 481 ggtgctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540  
 |||||  
 Db 481 ggtgctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540  
 QY 541 ctgagagctccagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 600  
 |||||  
 Db 541 ctgagagctccagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 600  
 QY 601 ctttcaacatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 660  
 |||||  
 Db 601 ctttcaacatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 660  
 QY 661 tctgagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
 |||||  
 Db 661 tctgagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
 QY 721 gggcagcctgagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 780  
 |||||  
 Db 721 gggcagcctgagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 780  
 QY 781 gcaagctcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 840  
 |||||  
 Db 781 gcaagctcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 840  
 QY 841 catcaacgcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900  
 |||||  
 Db 841 catcaacgcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900  
 QY 901 agagac 960  
 |||||  
 Db 901 agagac 960  
 QY 961 gct 1020  
 |||||  
 Db 961 gct 1020  
 QY 1021 gcaac 1080  
 |||||  
 Db 1021 gcaac 1080  
 QY 1081 ggcagctcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1140  
 |||||  
 Db 1081 ggcagctcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1140  
 QY 1141 atcagctcccgaggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200

Db 1141 atcagctcccgaggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200  
 |||||  
 QY 1201 ccacagctatagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1260  
 |||||  
 Db 1201 ccacagctatagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1260  
 QY 1261 ggaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
 |||||  
 Db 1261 ggaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
 QY 1321 cgaacagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380  
 |||||  
 Db 1321 cgaacagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380  
 QY 1381 catctccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440  
 |||||  
 Db 1381 catctccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440  
 QY 1441 tgaacacccgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1500  
 |||||  
 Db 1441 tgaacacccgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1500  
 QY 1501 ggggaaacacgagcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1560  
 |||||  
 Db 1501 ggggaaacacgagcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1560  
 QY 1561 gcaac 1620  
 |||||  
 Db 1561 gcaac 1620  
 QY 1621 gttgctctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1680  
 |||||  
 Db 1621 gttgctctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1680  
 QY 1681 caaac 1740  
 |||||  
 Db 1681 caaac 1740  
 QY 1741 acaaaagc 1800  
 |||||  
 Db 1741 acaaaagc 1800  
 QY 1801 ccaac 1860  
 |||||  
 Db 1801 ccaac 1860  
 QY 1861 cccgagaaatgataaagataactagaaactg 1894  
 |||||  
 Db 1861 cccgagaaatgataaagataactagaaactg 1894

RESULT 4  
 AAD07337  
 ID AAD07337 standard; cDNA; 1918 BP.  
 XX  
 AC AAD07337;  
 XX  
 DT 04-AUG-2001 (first entry)  
 XX  
 DE Human transcription factor, prost-ets cDNA.  
 XX  
 KW Human; transcription factor protein; PROST-Ets; cytosolic; gene therapy;  
 KM Vaccine; cancer; prostate; breast; ovary; benign prostatic hyperplasia;  
 KW antisense therapy; ribozyme therapy; therapeutic; cell proliferation; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 422..1429  
 FT /tag= a  
 FT /product= "Human PROST-Ets protein"  
 XX

PN WO200142472-A2.  
 XX 14-JUN-2001.  
 PD 29-NOV-2000; 2000WO-US42295.  
 XX 30-NOV-1999; 99US-0168182.  
 PR 21-NOV-2000; 2000US-0718159.  
 XX (SCHD ) SCHERING AG.  
 PA Harkins R, Lin RJ, Luke M, Montecclaro F, Parkes D, Parry G;  
 PI Steinhilber R, Van Heut PT, Xuan J;  
 XX MPI: 2001-381692/40.  
 DR P-SDB: AAE03556.  
 XX  
 PT New transcription factor polypeptides, designated PROST-Ets  
 PT polypeptides, useful for diagnostic, research and therapeutic  
 PT applications  
 PS Claim 8; Fig 1; 74pp; English.  
 XX  
 CC The invention relates to human transcription factor protein, designated  
 CC as PROST-Ets and its corresponding cDNA molecule. PROST-Ets proteins and  
 CC nucleic acids are useful for treating a disease-state in human patient,  
 CC such as prostate, breast and ovarian cancer which is associated with  
 CC inappropriate expression of PROST-Ets and where the patient is in need of  
 CC decreased levels or activity of PROST-Ets. PROST-Ets is useful for  
 CC diagnosing and treating diseases of cell proliferation, such as prostate  
 CC cancer and benign prostatic hyperplasia. Prost-ets nucleic acids are used  
 CC as DNA probes, as target for antisense and ribozyme therapy, as template  
 CC for producing antisense polynucleotides, and as diagnostic reagent for  
 CC research, biological, clinical and therapeutic purposes. Prost-ets DNA is  
 CC also used in gene therapy. PROST-Ets protein is useful for generating  
 CC antibodies, and for stimulating immune response in cells containing  
 CC PROST-Ets proteins. The present cDNA sequence encodes human transcription  
 CC factor, designated as PROST-Ets. Prost-ets cDNA was found in a cDNA clone  
 CC obtained from human prostate tissues.  
 CC  
 XX Sequence 1918 BP; 385 A; 656 C; 574 G; 303 T; 0 other:  
 SQ  
 Query Match 98.6%; Score 1868.4; DB 22; Length 1918;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 1878; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 gtctgactctccagacatctcgtcactcgtcgtgtccacactgcccacagacc 60  
 DB 7 gtctgactctccagacatctcgtcactcgtcgtgtccacactgcccacagacc 66  
 QY 61 agtctccaaagctgtgcaagctccctgcaagccctcgaagttggccttgcaagtg 120  
 DB 67 agtctccaaagctgtgcaagctccctgcaagccctcgaagttggccttgcaagtg 126  
 QY 121 ccaagcagcagcctggtggtggtgtagggactccctcagcagcagcctgtgacc 180  
 DB 127 ccaagcagcagcctggtggtggtgtagggactccctcagcagcagcctgtgacc 186  
 QY 181 gcaacacagcagccttgaggtggtgcaagcccccagtggtgcaaacctgtgctgctc 240  
 DB 187 tcagagggcagcccttgaggtggtgcaagcccccagtggtgcaaacctgtgctgctc 246  
 QY 241 gcaacacagccttggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 300  
 DB 247 gcaacacagccttggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 306  
 QY 301 cagtggcctcagctggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 360  
 DB 307 cagtggcctcagctggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 366  
 QY 361 ctccctgggacagcagcgttaacagacagcgcgcagcccaaaagcagcggtggtg 420  
 DB 367 ctccctgggacagcagcgttaacagacagcgcgcagcccaaaagcagcggtggtg 426

DB 367 ctccctgggacagcagcgttaacagacagcgcgcagcccaaaagcagcggtggtg 426  
 QY 421 cagcgcaagccgggtctgagcagcgtatcccccagccactctctgctgcccccgagac 480  
 DB 427 cagcgcaagccgggtctgagcagcgtatcccccagccactctctgctgcccccgagac 486  
 QY 481 ggtgtcggtgagcagcgttggtggtggtggtggtggtggtggtggtggtggtggtggtg 540  
 DB 487 ggtgtcggtgagcagcgttggtggtggtggtggtggtggtggtggtggtggtggtggtg 546  
 QY 541 ctgagttccagctccaccccgccagcagcagcagcagcagcagcagcagcagcagcagcagc 600  
 DB 547 ctgagttccagctccaccccgccagcagcagcagcagcagcagcagcagcagcagcagcagc 606  
 QY 601 ctgtgacatgtgtgacacttgagcagcagcagcagcagcagcagcagcagcagcagcagc 660  
 DB 607 ctgtgacatgtgtgacacttgagcagcagcagcagcagcagcagcagcagcagcagcagc 666  
 QY 661 tcggagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
 DB 667 tcggagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 726  
 QY 721 ggcagccctgagcttggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 780  
 DB 727 ggcagccctgagcttggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 786  
 QY 781 gcatcactgtgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 840  
 DB 787 gcatcactgtgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 846  
 QY 841 catcacgcagatcccatgtgagctgagcagcagcagcagcagcagcagcagcagcagcagc 900  
 DB 847 catcacgcagatcccatgtgagctgagcagcagcagcagcagcagcagcagcagcagcagc 906  
 QY 901 agagacccaataccggctgccccccatggcagcagcagcagcagcagcagcagcagcagcagc 960  
 DB 907 agagacccaataccggctgccccccatggcagcagcagcagcagcagcagcagcagcagcagc 966  
 QY 961 gctgtgcgcactgtcagagagcagctccgcagcagcagcagcagcagcagcagcagcagcagc 1020  
 DB 967 gctgtgcgcactgtcagagagcagctccgcagcagcagcagcagcagcagcagcagcagcagc 1026  
 QY 1021 gcaagcccaactgtgacatctggaagtcagcggtggtggtggtggtggtggtggtggtggtg 1080  
 DB 1027 gcaagcccaactgtgacatctggaagtcagcggtggtggtggtggtggtggtggtggtggtg 1086  
 QY 1081 ggcgaattcaactgtgctcgacagtgagagagcagcagcagcagcagcagcagcagcagcagc 1140  
 DB 1087 ggcgaattcaactgtgctcgacagtgagagagcagcagcagcagcagcagcagcagcagcagc 1146  
 QY 1141 atcatgtctcggtgagcccatcactgtgagcttccctcaagagagttgctactcaagcc 1200  
 DB 1147 atcatgtctcggtgagcccatcactgtgagcttccctcaagagagttgctactcaagcc 1206  
 QY 1201 ccaagcagctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1260  
 DB 1207 ccaagcagctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1266  
 QY 1261 ggaactcaagcccaagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1320  
 DB 1267 ggaactcaagcccaagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1326  
 QY 1321 cgaacagctgagcgtctcactcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380  
 DB 1327 cgaacagctgagcgtctcactcagcagcagcagcagcagcagcagcagcagcagcagcagc 1386  
 QY 1381 catctccagcagcgtctcactcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440  
 DB 1387 catctccagcagcgtctcactcagcagcagcagcagcagcagcagcagcagcagcagcagc 1446  
 QY 1441 tgaacccgcctcagaggtcctctcctgtgctgctgctgctgctgctgctgctgctgctgctg 1500  
 DB 1447 tgaacccgcctcagaggtcctctcctgtgctgctgctgctgctgctgctgctgctgctgctg 1506



QY	721	gvgbaacccvgauctvgvgccvgvgvgctvgaccccttvgaaagvgaactvgctvgagcaggt	760
Db	2120	gvgcagacccctvgacttvgctgcccvgvgctvgaccccttvgaaagvgaactvgctvgagcaggt	2179
QY	781	gcbagtcacatvgtvgvtvgvgvgaaagtvgctaaagvgaactvgaaagccvgctvgcaagctvgctcaa	840
Db	2180	gcbagtcacatvgtvgvtvgvgvgaaagtvgctaaagvgaactvgaaagccvgctvgcaagctvgctcaa	2239
QY	841	catcacccvgacatccacatvgacctvgagcccccagacaaatgtvgaaagatvgctccctgtvgac	900
Db	2240	catcacccvgacatccacatvgacctvgagcccccagacaaatgtvgaaagatvgctccctgtvgac	2299
QY	901	agagacacaaataccvgctvgcccccatgtvgcaagvgccttccagagagctvgcvgvgcaagva	960
Db	2300	agagacacaaataccvgctvgcccccatgtvgcaagvgccttccagagagctvgcvgvgcaagva	2359
QY	961	gctvtvgvgccatvtvgvgaaagvgaagtctccvgccaaagvgctcvgccccctvgvgvtvgvgatvgct	1020
Db	2360	gctvtvgvgccatvtvgvgaaagvgaagtctccvgccaaagvgctcvgccccctvgvgvtvgvgatvgct	2419
QY	1021	gcaagvgccaaactvgaaacatctvgaaagtcaagcvggctccgtvgaaatgaagvgaaatccacatvg	1080
Db	2420	gcaagvgccaaactvgaaacatctvgaaagtcaagcvggctccgtvgaaatgaagvgaaatccacatvg	2479
QY	1081	gvgcattccacatactvgctcctcgacacagvtgaagvaagctvgagcccaagccagvgctvgagctc	1140
Db	2480	gvgcattccacatactvgctcctcgacacagvtgaagvaagctvgagcccaagccagvgctvgagctc	2539
QY	1141	atcatvgctccvgvgaaagccacatccacactvgvgaaagtcttcccaagaaagtgtgtgtacccaagcc	1200
Db	2540	atcatvgctccvgvgaaagccacatccacactvgvgaaagtcttcccaagaaagtgtgtgtacccaagcc	2599
QY	1201	ccaaagvgctatvgccvgcttcatatgtvggtctcaacaagaaagvggacatctccaaaatvga	1260
Db	2600	ccaaagvgctatvgccvgcttcatatgtvggtctcaacaagaaagvggacatctccaaaatvga	2659
QY	1261	ggaactcagcccaagvtvgcccgvgctvtvgvggacatccvgcaagaaacgctcccgccatvgaaactg	1320
Db	2660	ggaactcagcccaagvtvgcccgvgctvtvgvggacatccvgcaagaaacgctcccgccatvgaaactg	2719
QY	1321	cgaaacagctvgccvgctcacaatccvgccaaagtcttcaaaagaaagvgacatcccgaaagccaga	1380
Db	2720	cgaaacagctvgccvgctcacaatccvgccaaagtcttcaaaagaaagvgacatcccgaaagccaga	2779
QY	1381	catctcccaagvgccctcgtctacacaggttcgtvgcaaccccaatctgaagtvgctvgccaaagvgcc	1440
Db	2780	catctcccaagvgccctcgtctacacaggttcgtvgcaaccccaatctgaagtvgctvgccaaagvgcc	2839
QY	1441	tgaaaacccvgccctcaagvggacctctctccvgctcvgccctcgaagccagvgccctvgaaagt	1500
Db	2840	tgaaaacccvgccctcaagvggacctctctccvgctcvgccctcgaagccagvgccctvgaaagt	2899
QY	1501	ggvgaaaaaacvgvgcagvtgtgtctctvgccvgctctvgaaaccttccaaagvgccaaagvtcaagvgaaag	1560
Db	2900	ggvgaaaaaacvgvgcagvtgtgtctctvgccvgctctvgaaaccttccaaagvgccaaagvtcaagvgaaag	2959
QY	1561	gcaacccaactvgccccagvgggaataatvggtctctctvgvggcttccvggaaacaaatvgvggcagvg	1620
Db	2960	gcaacccaactvgccccagvgggaataatvggtctctctvgvggcttccvggaaacaaatvgvggcagvg	3019
QY	1621	gtvgcttccctcctcagvgcccaagctvgctccctctvgaaagaaagaaagvgaagvgvgctcc	1680
Db	3020	gtvgcttccctcctcagvgcccaagctvgctccctctvgaaagaaagaaagvgaagvgvgctcc	3079
QY	1681	caaacacctvgctctvgaaacccaagcatcttccaaagvgaagvgccttcaaaagvggcagvtgaactcg	1740
Db	3080	caaacacctvgctctvgaaacccaagcatcttccaaagvgaagvgccttcaaaagvggcagvtgaactcg	3139
QY	1741	acaaagvgccaaagvgcagvtccaagvgcctctctctgtctcaatcccccctvgctcccatcttvgca	1800
Db	3140	acaaagvgccaaagvgcagvtccaagvgcctctctctgtctcaatcccccctvgctcccatcttvgca	3199
QY	1801	ccaaacccvgaatvgvgcagvggaagaaatctvgcaacccctcatggttvgcagccagvgaaagtgc	1860

Db	Accession	Length	Score	DB	Length
Db	3200 ccaacccggaatggtcagaggaagacacccggaatggtcagaggaagtc	3259	98.3%	DB 22	1905
Qy	1861 ccgaggaatgataaataagatactagagaactg	1894	99.2%	Pred. No. 0	
Db	3260 ccgaggaatgataatagatactagagaactg	3293			
RESULT	6				
AAH42468					
ID	AAH42468 standard; cDNA: 1905 BP.				
XX					
AC	AAH42468;				
DT	01-OCT-2001 (first entry)				
DE					
XX					
XX	Nucleotide sequence of human prostate associated Ets protein (PRAEP).				
XX					
XX	Human; prostate associated Ets protein; PRAEP; immunological disorder;				
KM	cell proliferation disorder; reproductive disorder; psoriasis; diabetes				
KM	developmental disorder; cancer; arteriosclerosis; multiple sclerosis;				
KW	acquired immunodeficiency syndrome; AIDS; polycystic ovary syndrome;				
XX	ectopic pregnancy; gonadal dysgenesis; achondroplastic dwarfism; ss.				
XX					
OS	Homo sapiens.				
FH	Key	Location/Qualifiers			
FT	CDS	430..1437			
FT		/*tag= a			
FT		/product= "prostate associated Ets protein (PRAEP)"			
XX					
XX	US6265565-B1.				
XX					
PD	24-JUL-2001.				
XX					
XX	03-APR-1998;	98US-0055113.			
PF					
XX	03-APR-1998;	98US-0055113.			
PA					
XX	(INCY-) INCYTE GENOMICS INC.				
PI	Bandman O, Guegler KJ, Lal P, Corley NC;				
XX					
DR	WPI: 2001-463945/50.				
XX	P-PSDB; AAG63017.				
PT					
PT	Polynucleotides encoding human prostate associated Ets proteins, useful				
PT	for preventing, diagnosing and treating, e.g. cancer, diabetes,				
XX	psoriasis and dwarfism -				
PS	Claim 3; Fig 1A-E; 31pp; English.				
XX					
XX					
CC	The present sequence encodes a human prostate associated Ets protein				
CC	(PRAEP). PRAEP may be used in the prevention, treatment and diagnosis				
CC	of diseases associated with inappropriate PRAEP expression such as				
CC	cell proliferation, immunological, reproductive and developmental				
CC	disorders. The PRAEP polypeptides may be used as antigens in the				
CC	production of antibodies against PRAEP and in assays to identify				
CC	modulators (agonists and antagonists) of PRAEP expression and activity.				
CC	The anti-PRAEP antibodies and PRAEP antagonists may also be used to				
CC	down regulate PRAEP expression and activity. The anti-PRAEP antibodies				
CC	and antagonists may be used to prevent and treat cell proliferation				
CC	disorders (e.g. cancer, arteriosclerosis and psoriasis), immunological				
CC	disorders (e.g. acquired immunodeficiency syndrome (AIDS), multiple				
CC	sclerosis and diabetes), reproductive disorders (e.g. polycystic ovary				
CC	syndrome and ectopic pregnancy) and disorders of growth and development				
XX	(e.g. gonadal dysgenesis and achondroplastic dwarfism).				
XX					
XX	Sequence 1905 BP; 365 A; 661 C; 576 G; 303 T; 0 other;				
Query Match	98.3%;	Score 1861.4;	DB 22;	Length 1905;	
Best Local Similarity	99.2%;	Pred. No. 0;			



	Matches 18/1;	Conservative	0;	Mismatches	16;	Indels	0;	Gaps	0;
QY	1	gtctgagcttcctcccaagacacatcttcctgacatctgcgtgtccatccatctgcccacagacc	60						
Db	15	gtctgagcttcctcccaagacacatcttcctgacatctgcgtgtccatccatctgcccacagacc	74						
QY	61	agctccccaagcctgctgcgcagctccctctgcagaagccctcaggtctgggctcttgccacggtg	120						
Db	75	agctccccaagcctgctgcgcagctccctctgcagaagccctcaggtctgggctcttgccacggtg	134						
QY	121	ccagccagagcagccctctggggttgaaaggacatccctacaaagagacagccaccccttgagacc	180						
Db	135	ccagccagagcagccctctggggttgaaaggacatccctacaaagagacagccaccccttgagacc	194						
QY	181	ggccacagccacccctctgagggctggccagagcccccagctggccaaactgagctgcctct	240						
Db	195	tcagagagccacccctctgagggctggccagagcccccagctggccaaactgagctgcctct	254						
QY	241	ggccacagccctctctggccctggctcccggtctccgctctggccccccaaatctcctgggtctggagaccg	300						
Db	255	ggccacagccctctctggccctggctcccggtctccgctctggccccccaaatctcctgggtctggagaccg	314						
QY	301	caagtggcctcagagctggccacacatctctcccgccctctggagattggcacctggccagagacag	360						
Db	315	caagtggcctcagagctggccacacatctctcccgccctctggagattggcacctggagacag	374						
QY	361	cttccctggggccacacagacagctaacagacacacacgcccagcccaaaacagacggcatggg	420						
Db	375	cttccctggggccacacagacagctaacagacacacacgcccagcccaaaacagacggcatggg	434						
QY	421	caagccagcagcccggtgctctagcagcgttatcccccagccacatccctgtgtcccccgcgacaa	480						
Db	435	caagccagcagcccggtgctctagcagcgttatcccccagccacatccctgtgtcccccgcgacaa	494						
QY	481	ggctgtcgcgggacacagcctcttgaaagaaagccggccagccggggccagctgggtctccgaagaaagggaa	540						
Db	495	ggctgtcgcgggacacagcctcttgaaagaaagccggccagccggggccagctgggtctccgaagaaagggaa	554						
QY	541	cttgagagctcccaagctccaccccgccacagccccgaagagagcctgtccgcctcttgcaactctcta	600						
Db	555	cttgagagctcccaagctccaccccgccacagccccgaagagagcctgtccgcctcttgcaactctctcta	614						
QY	601	ctttgacacatgcttaccctctgaaaggaaacagcagctgggtcagccaaagccctctgggcccagcag	660						
Db	615	ctttgacacatgcttaccctctgaaaggaaacagcagctgggtcagccaaagccctctgggcccagcag	674						
QY	661	tcgggaagagacacacctgacacacaccttgaaagaaagctggcccggtctacttgaaagccaaagcccagc	720						
Db	675	tcgggaagagacacacctgacacacaccttgaaagaaagctggcccggtctacttgaaagccaaagcccagc	734						
QY	721	ggggacccctggaacttggttgcccgccgggcctggaactcttgaaagaaagacatgcctgctgaagcagt	780						
Db	735	ggggacccctggaacttggttgcccgccgggcctggaactcttgaaagaaagacatgcctgctgaagcagt	794						
QY	781	ggaaatccacatggttggttgggcgaaagtgtctaaagaaacatctgaagacacggccctggcgaagctgctcaca	840						
Db	795	ggaaatccacatggttggttgggcgaaagtgtctaaagaaacatctgaagacacggccctggcgaagctgctcaca	854						
QY	841	catcacccgcagatcccatctgagacttggaagcccccagacatgttgaaagaaagtggctcctgttgac	900						
Db	855	catcacccgcagatcccatctgagacttggaagcccccagacatgttgaaagaaagtggctcctgttgac	914						
QY	901	agagacacaaataccggctggccccccatctgggcaaaagccttccagagacgtctggggggcagaagga	960						
Db	915	agagacacaaataccggctggccccccatctgggcaaaagccttccagagacgtctggggggcagaagga	974						
QY	961	ggcttgagcagcatgtctggaagagagagatctccggccagagacgtccggccctctgggttggaatgtgtct	1020						
Db	975	ggcttgagcagcatgtctggaagagagagatctccggccagagacgtccggccctctgggttggaatgtgtct	1034						
QY	1021	ggacagccacactggagacatctggaagtccagcggcctggatggaagagcggagcttaccctgg	1080						
Db	1035	ggacagccacactggagacatctggaagtccagcggcctggatggaagagcggagcttaccctgg	1094						

QY	1081	ggcgattcactctgtgtcgtcgaccaggtgagagagctgtgacccgacagcgaggtgagctc	1140
Db	1095	ggcgattcactactgtgtcgtcgaccaggtgagagagctgtgacccgacagcgaggtgagctc	1154
QY	1141	atactgtcccgagcgccatccactgttggcagttctctcaagaaggttctactcaagcc	1200
Db	1155	atactgtcccgagcgccatccactgttggcagttctctcaagaaggttctactcaagcc	1214
QY	1201	ccacagctctatggccgcttccattaggttggctctcaacaaagagagagagctctcaaaatttga	1260
Db	1215	ccacagctatggccgcttccattaggttggctctcaacaaagagagagagagctctcaaaatttga	1274
QY	1261	ggactcagaccacaggttggcccggtctgttgggcatctccgaagaaacggtcccgcaattgactgt	1320
Db	1275	ggactcagaccacaggttggcccggtctgttgggcatctccgaagaaacggtcccgcaattgactgt	1334
QY	1321	cgacacagcttgagccgcttccatccgcacagttcttaagaagaaggagctatccctggaaagccaga	1380
Db	1335	cgacacagcttgagccgcttccatccgcacagttcttaagaagaaggagagctatccctggaaagccaga	1394
QY	1381	catctccacgcgccttcgtctactacagattcgttgacacccacatctgaatgtccttggcccgagggcc	1440
Db	1395	catctccacgcgccttcgtctactacagattcgttgacacccacatctgaatgtccttggcccgagggcc	1454
QY	1441	tggaaacccgctctcaaggggcctctctctctgtccttgccttgcctcaagccagccctgagatg	1500
Db	1455	tggaaacccgctctcaaggggcctctctctctgtccttgccttgcctcaagccagcccttgaagatg	1514
QY	1501	ggggaaagcggcaggttgcctctgtctctgtacccctccagagcccaagagttcagagagag	1560
Db	1515	ggggaaagcggcaggttgcctctgtctctgtacccctccagagcccaagagttcagagagag	1574
QY	1561	gcaaccaactgcccacaggggagataggttcctcttggggccttcgggaacatgtgggcaggg	1620
Db	1575	gcaaccaactgcccacaggggagataggttcctcttggggccttcgggaacatgtgggcaggg	1634
QY	1621	gtgtcttcctctcagggcccaagcttgtctccctcgaggagacagagggagagagaggtctctccc	1680
Db	1635	gtgtcttcctctcagggcccaagcttgtctccctcgaggagacagagggagagagaggtctctccc	1694
QY	1681	caaacactgctcttgagccccaagcatcttcacagagacagagctctaaagaaaggcgagttgactcg	1740
Db	1695	caaacactgctcttgagccccaagcatcttcacagagacagagctctaaagaaaggcgagttgactcg	1754
QY	1741	acaaaggccaagagcagttccagagcctctctctgtcctcatcccccctgtcctccattctgca	1800
Db	1755	acaaaggccaagagcagttccagagcctctctctgtcctcatcccccctgtcctccattctgca	1814
QY	1801	ccaacactgtgcagtgtgcagggagagacatctcacccctccagtttgggagacagagagttgc	1860
Db	1815	ccaacactgtgcagtgtgcagggagagacatctcacccctccagtttgggagacagagagttgc	1874
QY	1861	cccggaatgtgataataaagatactag	1887
Db	1875	cccggaatgtgataataaagatactag	1901
RESULT 7			
AAFP1828			
ID	AAFP21828 standard; DNA; 1087 BP.		
XX	AAFP21828:		
AC			
XX			
DT	27-MAR-2001 (first entry)		
DE	Human breast and ovarian cancer associated antigen gene SHQ ID 215.		
XX			
KW	Human: breast cancer; ovarian cancer; cytostatic; immunosuppressive; neurotrophic; neuroprotective; antiviral; antiallergic; hepatoprotective; antidiabetic; antiinflammatory; antitumor; vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic; cardiant; immune disorder; Addison's disease; allergy; autoimmune haemolytic anaemia;		

Query Match	53.8%	Score 1018.6	DB 21	Length 1087
Best Local Similarity	97.0%	Pred. No. 4.8e-205		
Matches 1055	Conservative	3	Mismatches 28	Indels 2
			Gaps	2
QY	541	ctggagatcccaagtcacccgcaagccgagcaagggcgtctccgcttcgactctccta	600	
DB	2	ctggagatcccaagtcacccgcaagccgagcaagggcgtctccgcttcgactctccta	61	
QY	601	cttgcgcatgcgtctaccctggaggaagcagcctgggagcaagccctcggggccagcag	660	
DB	62	cttgcgcatgcgtctaccctggaggaagcagcctgggagcaagccctcggggccagcag	121	
QY	661	tcgggagggagccac	720	
DB	122	tcgggagggagccac	181	
QY	721	gggac	780	
DB	182	gggac	241	
QY	781	gcagcgcac	840	

Db	242	gcagcgcacgaatggtggtggtgcgaagtgtcccaagacatcgaagacatcgagccctgcgaagctgtcctca	301
QY	841	catcacccgaatattcccatctggaacttgagagagagccccaagaacattgtgcagaaagtgtgctctctgtgac	900
Db	302	catcaaccgacagatcccatctggaacttgagagagagccccaagaacattgtgcagaaagtgtgctctctgtgac	361
QY	901	agagacaccaataaccggtctgcgcgcccaatggtgcgaagccttcacagagactgtgcgtgcgaagaga	960
Db	362	agagacaccaataaccggtctgcgcgcccaatggtgcgaagccttcacagagactgtgcgtgcgaagaga	421
QY	961	gctgtgcgcacatgtgcgcgaagagagacagttctccgcgcgaagcgctccgcgccttggtgtggtgagatgtct	1020
Db	422	gctgtgtgcgcacatgtgcgcgaagagagacagttctccgcgcgaagcgctccgcgccttggtgtggtgagatgtct	481
QY	1021	gcacgcgcccaaccctgtgacatctctggaagatcgaagcgctctggaatgaagaagagcgagacttcacccgtg	1080
Db	482	gcacgcgcccaaccctgtgacatctctggaagatcgaagcgctctggaatgaagaagagcgagacttcacccgtg	541
QY	1081	ggcgatctcaactactgtgtctctgcgcacagctgagagagagctgtgagacgcagacgcagagctgtgactc	1140
Db	542	ggcgatctcaactactgtgtctctgcgcacagctgagagagagctgtgagacgcagacgcagagctgtgactc	601
QY	1141	atcatctgtcccggtgcgaagcccatctccacccgtgtgcgaagttctccccaagggaggttctacttcaaacgc	1200
Db	602	atcatctgtcccggtgcgaagcccatctccacccgtgtgcgaagttctccccaagggaggttctacttcaaacgc	661
QY	1201	ccaacagactatgtgcgcgcttcatctgaagctgtgcctcacaagaagagagagcgatctccaaataatgtga	1260
Db	662	ccaacagactatgtgcgcgcttcatctgaagctgtgcctcacaagaagagagagcgatctccaaataatgtga	721
QY	1261	gggaactcagcccagatgtgagcccggtctgtggtggtgcattcgcgaagaacgctcccgccaatgaactgt	1320
Db	722	gggaactcagcccagatgtgagcccggtctgtggtggtgcattcgcgaagaacgctcccgccaatgaactgt	781
QY	1321	cgacaagaagcttgagccgcgcctccatctccgcgaagttcttaacaagaagggcgatcatctccgaagaaccga	1380
Db	782	cgacaagaagcttgagccgcgcctccatctccgcgaagttcttaacaagaagggcgatcatctccgaagaaccga	841
QY	1381	catctcccaagagagctctgtctacacagattctgtgacaccccatctgaatgtgctgtgcgcacagggcc	1440
Db	842	catctcccaagagagctctgtctacacagattctgtgacaccccatctgaatgtgctgtgcgcacagggcc	901
QY	1441	tgaaaccgcgcctcctcaagggagcctctctctctgtgcctgtgcctcctcagccacagagccctvtgagatg	1500
Db	902	tgaaaccgcgcctcctcaagggagcctctctctctgtgcctgtgcctcctcagccacagagccctvtgagatg	961
QY	1501	gggggaaaaacgggcagctgtgtcctctgcgtctgtgaccttccacagagccccaagggccaaggtcaagagag	1560
Db	962	gggggaaaaac-ggcagctgtgtcctctgcgtcgtcgtcgtcctt-cagagcccaaggtccaagagagag	1019
QY	1561	gcaacccaacttgcacacagggggagataatggtctctctctgtgagccttcgtgagacattgtggtgcacag	1620
Db	1020	gcaacccaacttgcacacagggggagataatggtctctctctgtgagccttcgtgagacattgtggtgcacag	1079
QY	1621	gtgtgtctcc 1628	
Db	1080	gggcttcc 1087	
RESULT 8			
AAH47006	ID	AAH47006 standard; cDNA; 1704 BP.	
XX	AC	AAH47006;	
XX	DT	29-Oct-2001 (first entry)	
DE	Mouse ETS suppressor factor (esf) protein encoding cDNA.		
XX	ETS suppressor factor; esf; apoptosis; cancer; cytostatic; mouse;		
KM	growth-suppressor; transcription factor; gene-therapy; ss.		
XX	Mus sp.		



```
XX Key Location/Qualifiers
FH CDS 362..1339
FT /*tag= a
FT /product= "mouse esf"
PN MO200157186-A2.
XX 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US03640.
XX 04-FEB-2000; 2000US-0180244.
XX (MUSC-) MUSC FOUND RES DEV.
XX Watson DK:
XX WPI; 2001-465703/50.
XX P-PSDB; AAB85610.
XX Fragment of a human EMS suppressor factor protein for inhibiting the
XX growth or proliferation of and/or inducing apoptosis of cancer cells,
XX especially prostate, breast, colon or cervical cancer cells for
XX treatment.
XX Examples: Fig 1C; 94pp; English.
XX The invention relates to methods for treating or preventing cancer by
XX modifying the expression of EMS suppressor factor (esf) protein gene or
XX the activity of the gene product. The esf protein or its modified form
XX are useful for inhibiting the growth or proliferation of and/or inducing
XX apoptosis of cancer cells, especially prostate, breast, colon or cervical
XX cancer cells. They are useful for treating cancer in a subject when
XX introduced in to the vicinity of the cancer in a subject. The methods
XX provided are useful for inhibiting the specific degradation of esf
XX protein in a cell; for diagnosing a preneoplastic or neoplastic condition
XX in a human by e.g. PCR and for identifying a compound that inhibits the
XX degradation of esf protein in cancer cells. The present sequence
XX represents a mouse esf protein encoding cDNA.
XX Sequence 1704 BP; 369 A; 527 C; 491 G; 317 T; 0 other;
Query Match 39.0%; Score 738.4; DB 22; Length 1704;
Best Local Similarity 73.3%; Pred. No. 4,4e-146;
Matches 1061; Conservative 0; Mismatches 336; Indels 51; Gaps 7;
QY 38 tgtccacactgtccacacagcagctctccaaagtgtgcaagctcctctgcaagccccc 97
DB 1 tgtccgctctgtcctccacacactagaccacccagccc--gctcgtgccccggtgagaaacc 58
QY 98 tcaagttggcccttgccacaggtgtccacagcagccctgagctggtggttagggagctccc 157
DB 59 ccaagttgtgacctgtctcgtgacagcagtggtgacctgtcctcgtacacaggggacctgccc 118
QY 158 tacagcagcagcagcccttagacacccacaccccttgaagggtggtccagggcccccag 217
DB 119 tatagacttgacgctctgagc-----tcagactcacaactcaagggcagaagagctctg 174
QY 218 tggcacaactgagtgtcctctcgtccacacagccctgctggtggttccgcgtggtccccc 277
DB 175 tggccacactaagacgcccctctgtccccagcctgtgctgcccacatgat-----gtctg 228
QY 278 ccaagtgctctgtgtagacagcagcagtggtcagctggtcccaacacacttcccggtccccc 337
DB 229 actgagacacacagctgagccc-----tgagctgctctgcctgcacactgtccctctcgttcc 283
QY 338 ggaagttgacctgagcagcagacagctccctggtgacacacagcagcagctaaagacacagccgc 397
DB 284 gaggttgctctgtccgagagcagcagactctctcgaagcagggcggtcctaagcgaagcagcc 343
QY 398 agcccaaacagcagcagcagctggtgagcagcagcccggtgtctgagcagctatcccccagc 457
```

```
DB 344 ccaagctccacacgcagcagctggtgcaagtgtccagcccaagccttgaaacagtgctcccccgtc 403
QY 458 caactcctgtgtcccccacagcagtgctgcgagcagagcttgagaaagcgcagcggg 517
DB 404 tgcctgtcactgttccacagatgtgcaaccacagcagagcagcagcagcagcagcagc 463
QY 518 gcaagtgggtctccagagcagcagctgtggttccagctccacccgcagcccgagcagggc 577
DB 464 gcaatggcccttgaaagcagagaaatgagtgctcagtcacacccgcaccccttgagcagggc 523
QY 578 ctgtccgctctcagacctctcactactgaca tgtctgtaaccttgagagacagcagctgggca 637
DB 524 ctgtcctcttctactactctctactacttaaca---tgtatccgcagcagtagcagctggtc 580
QY 638 gccaaagcccttggtggtccagcagcagtcggagagagccacactgacacagcctgagtcgccg 657
DB 581 gccaaagtcctccgaaagcccggtgcagggagagcaccacccgcagcagcagctgtccccc 640
QY 698 gtcaattgacagcagcagcccccagcagcagcctgtgactgtgtgcctggcggtgtgacctgtg 757
DB 641 gtcatgtgacagcagcagcctctctgtgagc-----agctgtg 673
QY 758 gagagcagcctgcgttgagcaggtgtgcaatgcattgtgtgtggtgagagtgctcgaagacatc 817
DB 674 gatagcactcgtctagagcaggtgtcaatgcattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 733
QY 818 ggaagcgccttgaaagctgtctcaacatccacacacacacacacacacacacacacacacac 877
DB 734 gagagcgccttgaaagctgtctcaacatccacacacacacacacacacacacacacacacac 793
QY 878 gtccaaagctgtgtctctgttgagacagagacacacacacacacacacacacacacacacac 937
DB 794 gtccaaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 853
QY 938 ttcccaagagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 997
DB 854 ttcccaagagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 913
QY 998 tgcgccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1057
DB 914 gcaacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 973
QY 1058 atgaaagcgcgacttcaactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1117
DB 974 atgaaagcgcgacttcaactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1033
QY 1118 tggaccacagcagagtgagactcaatcagctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1177
DB 1034 tggacgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1093
QY 1178 ctcaagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1237
DB 1094 ctgaaagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1153
QY 1238 ggaagagcagcttcaaatgtgaagacacacacacacacacacacacacacacacacacacacac 1297
DB 1154 ggaagagcagcttcaaatgtgaagacacacacacacacacacacacacacacacacacacacac 1213
QY 1298 aagaacgcgtcccgccatgaactgtgacaagagctgagcgcgtccatccgcagcttcaacaag 1357
DB 1214 aagaacgcgtcccgccatgaactgtgacaagagctgagcgcgtccatccgcagcttcaacaag 1273
QY 1358 aagggcatcatccgaaagcagacatctcccaagcgcgtctcctacccagttgtgtcaacccc 1417
DB 1274 aagggcatcatctgtcaaacccagacatctcccaagcgcgtctcctacccagttgtgtcaaccca 1333
QY 1418 atctgagtgcc-----tggccaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1473
DB 1334 gctctgagagcacaagagacacagagcgtctcaaacacctgtcccgacagcagcagcagcagcagc 1393
QY 1474 gccctgtgcc 1481
DB 111 | | |
```

Db 1394 gcctggtc 1401

RESULT 9  
AAH30613 standard: cDNA; 392 BP.

XX ID AAH30613  
AC AAH30613;  
XX  
XX 27-JUL-2001 (first entry)  
XX  
XX Human colon cancer cell line Km12L4-A cDNA library derived sequence #547.  
DE  
XX  
XX Human: diagnosis: colon cancer; cancer; malignant; chromosome mapping;  
KW detection; colon cancer cell line Km12L4-A; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200018916-A2.  
XX  
XX 06-APR-2000.  
XX  
XX 23-SEP-1999; 99WO-US22226.  
XX  
XX 28-SEP-1998; 98US-0102161.  
XX 28-SEP-1998; 98US-0102180.  
XX 29-SEP-1998; 98US-0102380.  
XX 08-OCT-1998; 98US-0103815.  
XX 27-OCT-1998; 98US-0105877.  
XX  
XX (CHIR ) CHIRON CORP.  
XX (HYSE-) HYSEQ INC.  
XX  
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;  
PI Lamson G, Drmanac R, Dickson M, Drmanac S, Labat I;  
PI Leshkowitz D, Kila D, Garcia V, Jones LW, Stache-Crain B;  
XX  
XX WPI: 2000-293155/25.  
XX  
XX Polynucleotide library comprising 1079 defined sequences, useful in  
PT the form of an array to detect cancer or susceptibility to cancer -  
XX  
XX Claim 1: Page 340; 502pp; English.

The present invention describes a library of polynucleotides comprising 1079 nucleotide sequences (given in AAH3067 to AAH31145). Also described are: (1) an isolated polynucleotide (I) having at least 90% identity to one of the 1079 sequences; (2) a recombinant host cell containing (I); (3) an isolated polypeptide (II) encoded by (I); (4) an antibody that specifically binds to (II); (5) a vector comprising (I); and (6) a method of detecting differentially expressed genes correlated with a cancerous state of a mammalian cell comprising detecting a gene product encoded by 65 of the 1079 sequences given in the specification. The polynucleotides are used to monitor patients having (or susceptible) to cancer to detect potentially malignant events at a molecular level, before they are detectable at a gross morphological level. The polynucleotides are also useful for monitoring the efficacy of various therapies and preventive interventions. Polynucleotide probes based on the disclosed sequences are useful for chromosome mapping and detection of transcription levels. The 1079 polynucleotide sequences were derived from a human colon cancer cell line Km12L4-A cDNA library.

XX  
XX  
XX Sequence 392 BP; 82 A; 119 C; 126 G; 65 T; 0 other;

Query Match 20.6%; Score 391; DB 21; Length 392;  
Best Local Similarity 100.0%; Freq. NO. 3.9e-73;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 799 cgaagtcctcaagacatcgaagcgcgtcgaagtcctcaacatccgcagatcccat 858  
|||  
XX 1 cgaagtcctcaagacatcgaagcgcgtcgaagtcctcaacatccgcagatcccat 60

QY 859 ggaactgagagcccaagcaatgtcagaagtggctcctgtgagacagaccacaataccgct 918  
|||  
Db 61 ggaactgagagcccaagcaatgtcagaagtggctcctgtgagacagaccacaataccgct 120  
|||

QY 919 gcccccattggcaagagccttccagagagctgagcggaagagctgtgccaatgtcga 978  
|||  
Db 121 gcccccattggcaagagccttccagagagctgagcggaagagctgtgccaatgtcga 180  
|||

QY 979 ggaagcattccccaagcgtcgcctcctgggtggatgtgtctgcaagcccaactggact 1038  
|||  
Db 181 ggaagcattccccaagcgtcgcctcctgggtggatgtgtctgcaagcccaactggact 240  
|||

QY 1039 ctggaatcagcgccctgagatgaagaagcgacttaccctggggcgacttaccctgtgc 1098  
|||  
Db 241 ctggaatcagcgccctgagatgaagaagcgacttaccctggggcgacttaccctgtgc 300  
|||

QY 1099 ctggaacagtgagagagctggaaccagcgaagtgagatcatcatgtctccggcagcc 1158  
|||  
Db 301 ctggaacagtgagagagctggaaccagcgaagtgagatcatcatgtctccggcagcc 360  
|||

QY 1159 catccactgtggaagtcctcctcaaggagtgtg 1189  
|||  
Db 361 catccactgtggaagtcctcctcaaggagtgtg 391  
|||

RESULT 10  
AAC77746  
ID AAC77746 standard: cDNA; 427 BP.

XX  
XX AAC77746;  
XX  
XX 08-FEB-2001 (first entry)  
XX  
XX Human cancer associated gene sequence SEQ ID NO:140.

XX  
XX Human: cancer associated gene; cancer antigen; detection; cancer;  
KW diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;  
KW antidiabetic; antihistaminic; antirheumatic; antiallergic; antiviral;  
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;  
KW dermatological; neuroprotective; thrombolytic; coagulant; noctropic;  
KW vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;  
KW immune disorder; hematopoietic cell disorder; autoimmune disorder;  
KW allergic reaction; graft versus host disease; organ rejection;  
KW hemostatic; thrombolytic; cardiovascular disorder; infection;  
KW neurological disease; drug screening; ss.

XX  
XX Homo sapiens.  
XX  
XX WO20005350-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000WO-US05882.  
XX  
XX 12-MAR-1999; 99US-0124270.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
PI WPI: 2000-587533/55.  
XX P-PSDB; AAB43537.  
XX  
XX Novel isolated nucleic acids comprising sequences encoding peptides  
PT useful for treating or diagnosing e.g. cancer -  
XX  
XX Claim 1: Page 720-721; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given  
CC in AAB43398 to AAB44239. The proteins can have activities based on the  
XX tissues and cells the genes are expressed in. Example of activities  
CC include: cytostatic; proliferative; vulnerable; immunomodulator;









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 18:17:31 ; Search time 3380.83 Seconds  
(without alignments)  
12120.248 Million cell updates/sec

Title: US-09-841-963a-1  
Perfect score: 1894  
Sequence: 1 gctgacttcctccagcac.....ataaagactagagaaactg 1894

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues  
Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

```

Pending Patents_NA_Main:
1: /cgn2_6/prodata/2/pna/PCTUS_COMB.seq.*
2: /cgn2_6/prodata/2/pna/US06_COMB.seq.*
3: /cgn2_6/prodata/2/pna/US07_COMB.seq.*
4: /cgn2_6/prodata/2/pna/US08_COMB.seq.*
5: /cgn2_6/prodata/2/pna/US081_COMB.seq.*
6: /cgn2_6/prodata/2/pna/US082_COMB.seq.*
7: /cgn2_6/prodata/2/pna/US083_COMB.seq.*
8: /cgn2_6/prodata/2/pna/US084_COMB.seq.*
9: /cgn2_6/prodata/2/pna/US085_COMB.seq.*
10: /cgn2_6/prodata/2/pna/US086_COMB.seq.*
11: /cgn2_6/prodata/2/pna/US087_COMB.seq.*
12: /cgn2_6/prodata/2/pna/US088_COMB.seq.*
13: /cgn2_6/prodata/2/pna/US089_COMB.seq.*
14: /cgn2_6/prodata/2/pna/US090_COMB.seq.*
15: /cgn2_6/prodata/2/pna/US091_COMB.seq.*
16: /cgn2_6/prodata/2/pna/US092_COMB.seq.*
17: /cgn2_6/prodata/2/pna/US093_COMB.seq.*
18: /cgn2_6/prodata/2/pna/US094_COMB.seq.*
19: /cgn2_6/prodata/2/pna/US095A_COMB.seq.*
20: /cgn2_6/prodata/2/pna/US095B_COMB.seq.*
21: /cgn2_6/prodata/2/pna/US095C_COMB.seq.*
22: /cgn2_6/prodata/2/pna/US096_COMB.seq.*
23: /cgn2_6/prodata/2/pna/US096A_COMB.seq.*
24: /cgn2_6/prodata/2/pna/US096B_COMB.seq.*
25: /cgn2_6/prodata/2/pna/US096C_COMB.seq.*
26: /cgn2_6/prodata/2/pna/US096D_COMB.seq.*
27: /cgn2_6/prodata/2/pna/US096E_COMB.seq.*
28: /cgn2_6/prodata/2/pna/US097A_COMB.seq.*
29: /cgn2_6/prodata/2/pna/US097B_COMB.seq.*
30: /cgn2_6/prodata/2/pna/US097C_COMB.seq.*
31: /cgn2_6/prodata/2/pna/US098A_COMB.seq.*
32: /cgn2_6/prodata/2/pna/US098B_COMB.seq.*
33: /cgn2_6/prodata/2/pna/US098C_COMB.seq.*
34: /cgn2_6/prodata/2/pna/US099A_COMB.seq.*
35: /cgn2_6/prodata/2/pna/US099B_COMB.seq.*
36: /cgn2_6/prodata/2/pna/US099C_COMB.seq.*
37: /cgn2_6/prodata/2/pna/US100_COMB.seq.*
38: /cgn2_6/prodata/2/pna/US101_COMB.seq.*
39: /cgn2_6/prodata/2/pna/US6000_COMB.seq.*
40: /cgn2_6/prodata/2/pna/US6001_COMB.seq.*
41: /cgn2_6/prodata/2/pna/US6002_COMB.seq.*
42: /cgn2_6/prodata/2/pna/US6003_COMB.seq.*
43: /cgn2_6/prodata/2/pna/US6004_COMB.seq.*

```

Result No.	Score	Query Match %	Length	ID	Description
1	1894	100.0	1894	1 PCT-US01-03640-1	Sequence 1, App11
2	1894	100.0	1894	32 US-09-841-963a-1	Sequence 1, App11
3	1870	98.7	1894	1 PCT-US99-17470-1	Sequence 1, App11
4	1870	98.7	1894	15 US-09-126-945-1	Sequence 1, App11
5	1870	98.7	1894	15 US-09-126-945A-1	Sequence 1, App11
6	1870	98.7	1894	15 US-09-126-945B-1	Sequence 1, App11
7	1870	98.7	1894	22 US-09-570-593-1	Sequence 1, App11
8	1868.4	98.6	1902	14 US-09-064-645-14	Sequence 14, App1
9	1868.4	98.6	1902	14 US-09-064-645-15	Sequence 15, App1
10	1868.4	98.6	1902	22 US-09-567-970-14	Sequence 14, App1
11	1868.4	98.6	1902	22 US-09-567-970-15	Sequence 15, App1
12	1868.4	98.6	1918	28 US-09-718-159-1	Sequence 15, App1
13	1865.2	98.3	3317	22 US-09-570-593-12	Sequence 1, App11
14	1861.4	98.3	1905	33 US-09-866-356-2	Sequence 12, App1
15	1837.8	97.0	1907	59 US-60-209-009-74	Sequence 2, App11
16	1824.4	96.3	1916	56 US-60-172-373-4734	Sequence 74, App1
17	1803	95.2	1841	12 US-08-845-995-12	Sequence 12, App1
18	1019.8	53.8	1087	1 PCT-US00-05881-215	Sequence 215, App
19	1019.8	53.8	1087	34 US-09-925-298-215	Sequence 215, App
20	790.6	41.7	1188	59 US-60-209-009-75	Sequence 75, App1
21	790.6	41.7	1188	71 US-60-324-185-16637	Sequence 16637, A
22	755.2	39.9	1701	75 US-60-360-207-11811	Sequence 11811, A
23	626.4	33.1	1704	1 PCT-US01-03640-3	Sequence 3, App11
24	626.4	33.1	1704	32 US-09-841-963a-1	Sequence 3, App11
25	484.8	25.6	792	14 US-09-016-434-840	Sequence 840, App
26	459.4	24.3	478	25 US-09-644-869-2777	Sequence 2777, App
27	459.4	24.3	478	25 US-09-649-163-2650	Sequence 2650, App
28	434.6	22.9	609	20 US-09-535-897-20568	Sequence 20568, A
29	425	22.4	450	19 US-09-528-409-87128	Sequence 87128, A
30	425	22.4	450	35 US-09-933-524-87128	Sequence 87128, A
31	425	22.4	450	35 US-09-933-524A-87128	Sequence 87128, A

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES





|||||  
Db 1501 ggggaaaagggcagtgctgtctgtctgtaccttcagagcccaagtgcaaggagg 1560  
QY 1561 gcaaccaatgccccaggggagataggtctctgtggtcttcgggacatcagggcagg 1620  
Db 1561 gcaaccaatgccccaggggagataggtctctgtggtcttcgggacatcagggcagg 1620  
QY 1621 gtgtctctctcctcagccccagctgtctcccttgaggagagaggagagaggtgtctcc 1680  
Db 1621 gtgtctctctcctcagccccagctgtctcccttgaggagagaggagagaggtgtctcc 1680  
QY 1681 caaacactgtcttgaccccccaagcatttcagagcagagctcacaagggcgagtgctcc 1740  
Db 1681 caaacactgtcttgaccccccaagcatttcagagcagagctcacaagggcgagtgctcc 1740  
QY 1741 acaaaagccacagcagctcagcctctctgtctccatcccttcgctcccatctgca 1800  
Db 1741 acaaaagccacagcagctcagcctctctgtctccatcccttcgctcccatctgca 1800  
QY 1801 ccaacactgtgcatgtgtcagagagacatctgcacccctcagttggtgcagccaagagtc 1860  
Db 1801 ccaacactgtgcatgtgtcagagagacatctgcacccctcagttggtgcagccaagagtc 1860  
QY 1861 cccgggaatggtataaagatactagagaactg 1894  
Db 1861 cccgggaatggtataaagatactagagaactg 1894

RESULT 2  
US-09-841-963A-1  
Sequence 1. Application US/09841963A  
GENERAL INFORMATION:  
APPLICANT: Watson, Dennis K.  
APPLICANT: Papas, Takis S. (Deceased)  
APPLICANT: Papas, Tula C. (Legal Representative)  
TITLE OF INVENTION: Methods and compositions for the diagnosis and treatment of cancer  
TITLE OF INVENTION: based on transcription factor ETS2  
FILE REFERENCE: 10545-015-999  
CURRENT APPLICATION NUMBER: US/09/841,963A  
CURRENT FILING DATE: 2001-04-25  
PRIOR APPLICATION NUMBER: PCT/US99/27805  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: 06/109,850  
PRIOR FILING DATE: 1998-11-25  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1  
LENGTH: 1894  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (416) ..(1423)  
US-09-841-963A-1

Query Match 100.0%; Score 1894; DB 32; Length 1894;  
Best Local Similarity 100.0%; Freq. No. 0;  
Matches 1894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtctgactctctccagcacatctctgcaactgtccgtgtccacactgtccacagacc 60  
Db 1 gtctgactctctccagcacatctctgcaactgtccgtgtccacactgtccacagacc 60  
QY 61 agctctccagctgtctgcagctccctgtcaagccctcaggttgggtgtccaggtg 120  
Db 61 agctctccagctgtctgcagctccctgtcaagccctcaggttgggtgtccaggtg 120  
QY 121 ccagcagcagcctgtggtgtggtgagggagctccctcagcagcagcagcagcctagacc 180  
Db 121 ccagcagcagcctgtggtgtggtgagggagctccctcagcagcagcagcagcctagacc 180  
QY 181 gccaccagcaccctcttgaggtgtgcagagcccccagltgccaacctgtgtcctct 240

|||||  
Db 181 gccaccagcaccctcttgaggtgtgcagagcccccagltgccaacctgtgtcctct 240  
QY 241 gccaccagcaccctcttgaggtgtgcagaggttccgcttggcccccaagatggtgtgagacc 300  
Db 241 gccaccagcaccctcttgaggtgtgcagaggttccgcttggcccccaagatggtgtgagacc 300  
QY 301 cagtggtctcagctgtgccccacactctctcccggtcccttgaggtgtgacatcacaagacc 360  
Db 301 cagtggtctcagctgtgccccacactctctcccggtcccttgaggtgtgacatcacaagacc 360  
QY 361 ctcccttggtcagcagcagctaaacagacacagccgcagcccaaacagcagcgcatgg 420  
Db 361 ctcccttggtcagcagcagctaaacagacacagccgcagcccaaacagcagcgcatgg 420  
QY 421 cagtggtcagcagcagcttgagcagcgtatcccccagcaccctctctgtcccccgacc 480  
Db 421 cagtggtcagcagcagcttgagcagcgtatcccccagcaccctctctgtcccccgacc 480  
QY 481 ggtgtcgtgagcaggtcttgagaaagcgtgcagcgtggtgtgtctcagagagaggg 540  
Db 481 ggtgtcgtgagcaggtcttgagaaagcgtgcagcgtggtgtgtctcagagagaggg 540  
QY 541 ctgaggttccagttccaccccgccagcagcagcgtgtctccgcttgcactctccta 600  
Db 541 ctgaggttccagttccaccccgccagcagcagcgtgtctccgcttgcactctccta 600  
QY 601 ctgtgacatgtgtacccctgagcagcagcgtgtgcagcagcagcctgtggcgagcg 660  
Db 601 ctgtgacatgtgtacccctgagcagcagcgtgtgcagcagcagcctgtggcgagcg 660  
QY 661 tcgggagagagcaccctgagcagcagcgtgagcgtgctccggtcatctgacagccaagcccg 720  
Db 661 tcgggagagagcaccctgagcagcagcgtgagcgtgctccggtcatctgacagccaagcccg 720  
QY 721 ggggacctgagctgtgtgccccggcggtgtgacctgtgagagcactcgtgtgagaggt 780  
Db 721 ggggacctgagctgtgtgccccggcggtgtgacctgtgagagcactcgtgtgagaggt 780  
QY 781 gcaatccatgtgtgtgtgagaggtgtcacaagagacatcagaaagcgtgtcagatgtc 840  
Db 781 gcaatccatgtgtgtgtgagaggtgtcacaagagacatcagaaagcgtgtcagatgtc 840  
QY 841 catcacgcagatctccatgtgactgtgagccccaagcaatgtgcagaagtgtcctgtgtgac 900  
Db 841 catcacgcagatctccatgtgactgtgagccccaagcaatgtgtgcagaagtgtcctgtgtgac 900  
QY 901 agagcaccatacagcgtgtgcccccatgtggcagaagccttcagagagctgtgcgggagga 960  
Db 901 agagcaccatacagcgtgtgcccccatgtggcagaagccttcagagagctgtgcgggagga 960  
QY 961 gctgtgcagcagctgtcagagagcagcttcgcagcagcgtctgcgtccctgtgtgtgtgt 1020  
Db 961 gctgtgcagcagctgtcagagagcagcttcgcagcagcgtctgcgtccctgtgtgtgtgt 1020  
QY 1021 gcaagcccaactgtgacatctggaagtcagcggtgtgagatgaagaagcggtcactgtg 1080  
Db 1021 gcaagcccaactgtgacatctggaagtcagcggtgtgagatgaagaagcggtcactgtg 1080  
QY 1081 ggcatttcaactgtgtcctgcagcagctgagagagagctgtgacccagcaggtgtgactc 1140  
Db 1081 ggcatttcaactgtgtcctgcagcagctgagagagagctgtgacccagcaggtgtgactc 1140  
QY 1141 atcatgtctcgggtgagcccatcactgtgtgaggttctcctaagaggtgtgtactcaagcc 1200  
Db 1141 atcatgtctcgggtgagcccatcactgtgtgaggttctcctaagaggtgtgtactcaagcc 1200  
QY 1201 ccagagctatgtgcggttcatatgagtggtcacaagaagagagagaggttctcctaagaa 1260  
Db 1201 ccagagctatgtgcggttcatatgagtggtcacaagaagagagagaggttctcctaagaa 1260  
QY 1261 ggaactcagcccaagtggtccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1320

Db 1261 ggaactaagcccaagctgagcccgctgtgaggcatccgaaagacgctcccgcatagactg 1320  
QY 1321 cgacaagctgaagcgcctccatcccgcaagctcttaacaagaaggcattatccggaagccaga 1380  
Db 1321 cgacaagctgaagcgcctccatcccgcaagctcttaacaagaaggcattatccggaagccaga 1380  
QY 1381 catctcccaagcgcctgcctbaacagctctgcaaccatctgaagctgctgagcccaaggcc 1440  
Db 1381 catctcccaagcgcctgcctbaacagctctgcaaccatctgaagctgctgagcccaaggcc 1440  
QY 1441 tgaaccgcgcctcaaggagggcctctctccgctgcctccatcagcccaaggccctgaagctg 1500  
Db 1441 tgaaccgcgcctcaaggagggcctctctccgctgcctccatcagcccaaggccctgaagctg 1500  
QY 1501 ggggaaaaacgagcgtgtgctctgctctgcaacttccagagcccaagagtcagggagagg 1560  
Db 1501 ggggaaaaacgagcgtgtgctctgctctgcaacttccagagcccaagagtcagggagagg 1560  
QY 1561 gcaacaaactgcccacgggggagatactggtctctctgagcccttcggagccatgaggagag 1620  
Db 1561 gcaacaaactgcccacgggggagatactggtctctctgagcccttcggagccatgaggagag 1620  
QY 1621 gtgtctctctctcaagcccaagcgtctccctctgagagacagaggagagagcgtctgctcc 1680  
Db 1621 gtgtctctctctcaagcccaagcgtctccctctgagagacagaggagagagcgtctgctcc 1680  
QY 1681 caaacactgctctgaccccaagcatttccagagcaagagcttaacagaggcagtgactcg 1740  
Db 1681 caaacactgctctgaccccaagcatttccagagcaagagcttaacagaggcagtgactcg 1740  
QY 1741 acaaaagcccaagcagctcagagcctctctctccatccctccctgctccatcttgcga 1800  
Db 1741 acaaaagcccaagcagctcagagcctctctctccatccctccctgctccatcttgcga 1800  
QY 1801 ccaacactgctgagctgagcagagagacatctgcaaccctcaagcttggagcagccagagtgcc 1860  
Db 1801 ccaacactgctgagctgagcagagagacatctgcaaccctcaagcttggagcagccagagtgcc 1860  
QY 1861 cccgggaatgataataaagatactagagaactg 1894  
Db 1861 cccgggaatgataataaagatactagagaactg 1894

RESULT 3  
PCT-US99-17470-1  
: Sequence 1, Application FC/TUS9917470  
: GENERAL INFORMATION:  
: APPLICANT: Human Genome Sciences, Inc.  
: APPLICANT: Beth Israel Deaconess Medical Center  
: APPLICANT: Liberman, Towla A.  
: APPLICANT: Oeligen, Joerg P.  
: APPLICANT: Kunsch, Charles A.  
: APPLICANT: Endress, Gregory A.  
: APPLICANT: Rosen, Craig A.  
: TITLE OF INVENTION: Prostate Derived Ets Factor  
: FILE REFERENCE: 1488.109PC01  
: CURRENT APPLICATION NUMBER: PCT/US99/17470  
: EARLIER FILING DATE: 1999-08-02  
: EARLIER APPLICATION NUMBER: US 09/126,945  
: EARLIER FILING DATE: 1998-07-31  
: NUMBER OF SEQ ID NOS: 15  
: SOFTWARE: Patentin Ver. 2.0  
: SEQ ID NO 1  
: LENGTH: 1894  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
PCT-US99-17470-1

Query Match 98.7%; Score 1870; DB 1; Length 1894;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1879; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 gtctgacttctctccagacattctctgactctgctgctgtgtccactgcccacagacc 60  
Db 1 gtctgacttctctccagacattctctgactctgctgctgtgtccactgcccacagacc 60  
QY 61 agtctccaaagcctgctgctcagctctccctgcaagccctcaggttgaggctctggcagagtg 120  
Db 61 agtctccaaagcctgctgctcagctctccctgcaagccctcaggttgaggctctggcagagtg 120  
QY 121 caagcagcgaacccctgagctgaggtgaaagagacatccctacagcagcagccctgaagcc 180  
Db 121 caagcagcgaacccctgagctgaggtgaaagagacatccctacagcagcagccctgaagcc 180  
QY 181 gccaacagcgaacccctgaaggtgagcagcgcctccagctgccaactgagtgctgctctc 240  
Db 181 tcaaggggcacaccccttgagggtgagcagcgcctccagctgccaactgagtgctgctctc 240  
QY 241 gccacacagccctgctgagcccttgctccgctgagcccccacatgctcgtcgtgaacagcc 300  
Db 241 gccacacagccctgctgagcccttgctccgctgagcccccacatgctcgtcgtgaacagcc 300  
QY 301 cagtgactcagctgacccacacactctcccgccctgagattgagcactgacagacag 360  
Db 301 cagtgactcagctgacccacacactctcccgccctgagattgagcactgacagacag 360  
QY 361 ctccctgggcaacagcagcgtacacagacacagccgccaagcccaacagcagctggag 420  
Db 361 ctccctgggcaacagcagcgtacacagacacagccgccaacagcagctggag 420  
QY 421 cagcgcacagcccggtgtcagacagctatcccccagccactctctgctgcccccgac 480  
Db 421 cagcgcacagcccggtgtcagacagctatcccccagccactctctgctgcccccgac 480  
QY 481 ggtgtcagcgaacagccttgagagaaagcagcagcgaggagctggtctcagaaagacagga 540  
Db 481 ggtgtcagcgaacagccttgagagaaagcagcagcgaggagctggtctcagaaagacagga 540  
QY 541 ctgaggtcccaagcttccaccccgccagcccgccagagcctgtctccctcttccctcta 600  
Db 541 ctgaggtcccaagcttccaccccgccagcccgccagagcctgtctccctcttccctcta 600  
QY 601 ctttacaatgctgtaaccttgagagacagcagctgagcagcccaagcccttgaggcacaag 660  
Db 601 ctttacaatgctgtaaccttgagagacagcagctgagcagcccaagcccttgaggcacaag 660  
QY 661 tcggagagagccacactgacccagctgagcagctgcccgtatctacagccaagcccaagc 720  
Db 661 tcggagagagccacactgagcagcagcctgagcagctgcccgtatctacagccaagcccaagc 720  
QY 721 gggcaccctgagctgtgtgcccggcgagctgacacttgagagagacactcgtgagcaggt 780  
Db 721 gggcaccctgagctgtgtgcccggcgagctgacacttgagagagacactcgtgagcaggt 780  
QY 781 gcagtccatgtgtgtggcggaagtgtctcaagagacaacagacgagcctgcaagctgctcaa 840  
Db 781 gcagtccatgtgtgtggcggaagtgtctcaagagacaacagacgagcctgcaagctgctcaa 840  
QY 841 catcacccgaagatcccaatgagctgagccccaagcaatgtgagaagtgctctctgtgagc 900  
Db 841 catcacccgaagatcccaatgagctgagccccaagcaatgtgagaagtgctctctgtgagc 900  
QY 901 agagacacaaatccggctgagcccccacatgggcaagagccttccagagagctgaggccaagga 960  
Db 901 agagacacaaatccggctgagcccccacatgggcaagagccttccagagagctgaggccaagga 960  
QY 961 gctgtgcgcacatgtctgagagagcaatctccgcagagcgtcgcctctgagtgagtgact 1020  
Db 961 gctgtgcgcacatgtctgagagagcaatctccgcagagcgtcgcctctgagtgagtgact 1020  
QY 1021 gcaacgccaacttgaacatctggaagtcagcgagcttgaatgaagaagcggacttcaactg 1080  
Db 1021 gcaacgccaacttgaacatctggaagtcagcgagcttgaatgaagaagcggacttcaactg 1080  
QY 1081 ggcgattcaactctgtgctctgagcagctgagagagagcctgagccgacagcagagtgagctc 1140





QY 841 catcacccgagatcccatgactgagagccccagcaatgtcaggaagtgtgctctgtgac 900  
|||||  
Db 841 catcacccgagatcccatgactgagagccccagcaatgtcaggaagtgtgctctgtgac 900  
QY 901 agagcaaccaatacagctgagcccccctctgagcaagagcttcccaagagctgagcaagga 960  
|||||  
Db 901 agagcaaccaatacagctgagcccccctctgagcaagagcttcccaagagctgagcaagga 960  
QY 961 gctgtgctgagcaatctgaggaagagcagttccgccaagcgtcccgctgagtgagtgatgtct 1020  
|||||  
Db 961 gctgtgctgagcaatctgaggaagagcagttccgccaagcgtcccgctgagtgagtgatgtct 1020  
QY 1021 ggcagcccaactgagcaactctgaggaatcagagagcctgaggaagagagagcttccactg 1080  
|||||  
Db 1021 ggcagcccaactgagcaactctgaggaatcagagagcctgaggaagagagagcttccactg 1080  
QY 1081 ggcagatcactactgtgctctgagccagtgagagagagctgagccagagagagtgagctc 1140  
|||||  
Db 1081 ggcagatcactactgtgctctgagccagtgagagagagctgagccagagagagtgagctc 1140  
QY 1141 atcatgtctccgggagagcccatcactgtgagcagcttccccaagagagtgctactcaagcc 1200  
|||||  
Db 1141 atcatgtctccgggagagcccatcactgtgagcagcttccccaagagagtgctactcaagcc 1200  
QY 1201 ccaacagctatgagccgctctcaatgagtgagtgagcaagagagagagagcttcccaaatgga 1260  
|||||  
Db 1201 ccaacagctatgagccgctctcaatgagtgagtgagcaagagagagagagcttcccaaatgga 1260  
QY 1261 ggaactcagcccaagtgagcccgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1320  
|||||  
Db 1261 ggaactcagcccaagtgagcccgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1320  
QY 1321 cgaacaagctgagccgctctcaatcagctgtctcaagaagagagagagagagagagagagagag 1380  
|||||  
Db 1321 cgaacaagctgagccgctctcaatcagctgtctcaagaagagagagagagagagagagagagag 1380  
QY 1381 catctccagagcgcctgctcaacagagctgtgcaaccccatctgagtgctgtgtgtgtgtgtgt 1440  
|||||  
Db 1381 catctccagagcgcctgctcaacagagctgtgcaaccccatctgagtgctgtgtgtgtgtgtgt 1440  
QY 1441 tgaacacccgctctcag 1500  
|||||  
Db 1441 tgaacacccgctctcag 1500  
QY 1501 ggggagaaaacgagagagtgagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560  
|||||  
Db 1501 ggggagaaaacgagagagtgagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560  
QY 1561 gcaacccaactgagcccaag 1620  
|||||  
Db 1561 gcaacccaactgagcccaag 1620  
QY 1621 gtgtctctctcag 1680  
|||||  
Db 1621 gtgtctctctcag 1680  
QY 1681 caaacacctgctctgagcccaag 1740  
|||||  
Db 1681 caaacacctgctctgagcccaag 1740  
QY 1741 acaaaagcccaag 1800  
|||||  
Db 1741 acaaaagcccaag 1800  
QY 1801 ccacacacctgag 1860  
|||||  
Db 1801 ccacacacctgag 1860  
QY 1861 cccgggag 1894  
|||||  
Db 1861 cccgggag 1894

RESULT 6  
US-09-126-945B-1  
; Sequence 1, Application US/09126945B  
; GENERAL INFORMATION:  
; APPLICANT: Libermann, Towia A.  
; APPLICANT: Oetting, Joerg P.  
; APPLICANT: Kunsch, Charles A.  
; APPLICANT: Endress, Gregory A.  
; TITLE OF INVENTION: Prostate Derived Ets Factor  
; FILE REFERENCE: 1488.1090000  
; CURRENT APPLICATION NUMBER: US/09/126,945B  
; CURRENT FILING DATE: 1998-07-31  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1894  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-126-945B-1

Query Match 98.7%; Score 1870; DB 15; Length 1894;  
Best Local Similarly 99.2%; Pred. No. 0;  
Matches 1879; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 1 gctcagactctcccaagacatctctgactctgagctgagctgagctgagctgagctgagctgagctg 60  
|||||  
Db 1 gctcagactctcccaagacatctctgactctgagctgagctgagctgagctgagctgagctgagctg 60  
QY 61 agctctccaagctgt 120  
|||||  
Db 61 agctctccaagctgt 120  
QY 121 ccagcagagcagccctgt 180  
|||||  
Db 121 ccagcagagcagccctgt 180  
QY 181 gccaacacccacccctctgagagtgag 240  
|||||  
Db 181 tcagagagagcaccctctgagagtgag 240  
QY 241 gccaacacccctgt 300  
|||||  
Db 241 gccaacacccctgt 300  
QY 301 cagtgagctcagctgt 360  
|||||  
Db 301 cagtgagctcagctgt 360  
QY 361 ctccctgt 420  
|||||  
Db 361 ctccctgt 420  
QY 421 cagcgcagagcccggt 480  
|||||  
Db 421 cagcgcagagcccggt 480  
QY 481 ggtgtcagagcagagctgt 540  
|||||  
Db 481 ggtgtcagagcagagctgt 540  
QY 541 ctgagagctccagctcagcccgag 600  
|||||  
Db 541 ctgagagctccagctcagcccgag 600  
QY 601 ctcttacaatgtctgataccctgtgag 660  
|||||  
Db 601 ctcttacaatgtctgataccctgtgag 660  
QY 661 tcgggagagagcacaactgagcagcctgagcagtgcccggtgatactgacagcagcagcagcagcagc 720  
|||||

Db 661 tcggagagagccacatgagagcctgagcagctgcccgtcatctgacagcccaagc 720  
QY 721 ggggacccctggactgtgcccggcggtgtaaccttgagagagacctcctgagagagt 780  
Db 721 ggggacccctggactgtgcccggcggtgtaaccttgagagagacctcctgagagagt 780  
QY 781 gcaatcctgagctgtgagagagctgctcaagagacatcgagagagcctgcaagctctca 840  
Db 781 gcaatcctgagctgtgagagagctgctcaagagacatcgagagagcctgcaagctctca 840  
QY 841 catcacccagatccatctgagagcctgagcccaagcaatgtgcaagagtgctctgtgac 900  
Db 841 catcacccagatccatctgagagcctgagcccaagcaatgtgcaagagtgctctgtgac 900  
QY 901 agagcccaataacagcctgcccctcctcctcctcctcctcctcctcctcctcctcctc 960  
Db 901 agagcccaataacagcctgcccctcctcctcctcctcctcctcctcctcctcctcctc 960  
QY 961 gctgtgagcagctgctgagagagcagctcctcctcctcctcctcctcctcctcctcct 1020  
Db 961 gctgtgagcagctgctgagagagcagctcctcctcctcctcctcctcctcctcctcct 1020  
QY 1021 gcaagcccaatgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 1080  
Db 1021 gcaagcccaatgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 1080  
QY 1081 ggcagctcaactgctgctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1140  
Db 1081 ggcagctcaactgctgctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1140  
QY 1141 atcatgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1200  
Db 1141 atcatgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1200  
QY 1201 ccaagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctcct 1260  
Db 1201 ccaagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctcct 1260  
QY 1261 ggaatcagccagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcct 1320  
Db 1261 ggaatcagccagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcct 1320  
QY 1321 cgaagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1380  
Db 1321 cgaagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1380  
QY 1381 catctcccaagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctc 1440  
Db 1381 catctcccaagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctc 1440  
QY 1441 tgaagcagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 1500  
Db 1441 tgaagcagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 1500  
QY 1501 ggggagagagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcct 1560  
Db 1501 ggggagagagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcct 1560  
QY 1561 gcaagcagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 1620  
Db 1561 gcaagcagcagctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 1620  
QY 1621 gctgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1680  
Db 1621 gctgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1680  
QY 1681 caaacactgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1740  
Db 1681 caaacactgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1740  
QY 1741 acaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1800  
Db 1741 acaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1800

QY 1801 ccacactgagctgtgagagagacatctgacacccctcagcttgagcagcagagtgcc 1860  
Db 1801 ccacactgagctgtgagagagacatctgacacccctcagcttgagcagcagagtgcc 1860  
QY 1861 ccgagagatgataataagatactagagagctg 1894  
Db 1861 ccgagagatgataataagatactagagagctg 1894

RESULT 7  
US-09-570-593-1  
; Sequence 1, Application US/09570593  
; GENERAL INFORMATION:  
; APPLICANT: Kaufmann, Joerg  
; APPLICANT: Xin, Hong  
; APPLICANT: Harrowe, Greg  
; TITLE OF INVENTION: EXPRESSION OF EMS-DOMAIN PROTEINS IN  
; FILE REFERENCE: 2300-1556  
; CURRENT APPLICATION NUMBER: US/09/570,593  
; CURRENT FILING DATE: 2000-05-12  
; PRIOR APPLICATION NUMBER: 60/134,112  
; PRIOR FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1894  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (416)...(1423)  
; OTHER INFORMATION: Human JRETS.  
US-09-570-593-1

Query Match 98.7%; Score 1870; DB 22; Length 1894;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1879; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 gcttactctctccagacatctcctgacatctcctgctcctgctcctgctcctgctcctgctc 60  
Db 1 gcttactctctccagacatctcctgacatctcctgctcctgctcctgctcctgctcctgctc 60  
QY 61 agctctccagcctgtgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 120  
Db 61 agctctccagcctgtgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 120  
QY 121 ccagcagcagcctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 180  
Db 121 ccagcagcagcctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 180  
QY 181 gcaacagcagcctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 240  
Db 181 gcaacagcagcctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 240  
QY 241 gcaacagcagcctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 300  
Db 241 gcaacagcagcctgagcagctcctcctcctcctcctcctcctcctcctcctcctcctc 300  
QY 301 cagtgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 360  
Db 301 cagtgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 360  
QY 361 ctccttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
Db 361 ctccttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
QY 421 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480  
Db 421 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480



QY 481 ggtgtcgcgagcagcttgagaaagcgcgagcggtggtcgagctggtctcgagagacggtga 540  
|||||  
Db 481 ggtgtcgcgagcagcttgagaaagcgcgagcggtggtcgagctggtctcgagagacggtga 540  
QY 541 ctggagctcccaatccacccgacgacggagagagcgctgtccgctcttcgagccctctcta 600  
|||||  
Db 541 ctggagctcccaatccacccgacgacggagagagcgctgtccgctcttcgagccctctcta 600  
QY 601 ctttgacatgctgtacccctgagagacagcagctggtgacgacgacgacgacgacgacgacg 660  
|||||  
Db 601 ctttgacatgctgtacccctgagagacagcagctggtgacgacgacgacgacgacgacgacg 660  
QY 661 tggggagagagcagcctgacagcagctgagcagctgagcagctgagcagcagcagcagcagc 720  
|||||  
Db 661 tggggagagagcagcctgacagcagctgagcagctgagcagctgagcagcagcagcagcagc 720  
QY 721 gggcagccctgagctgt 780  
|||||  
Db 721 gggcagccctgagctgt 780  
QY 781 gcaagtcacatgt 840  
|||||  
Db 781 gcaagtcacatgt 840  
QY 841 catcacacgagatcccatgagctgagcagcagcagcagcagcagcagcagcagcagcagcagc 900  
|||||  
Db 841 catcacacgagatcccatgagctgagcagcagcagcagcagcagcagcagcagcagcagcagc 900  
QY 901 aagagcacaataccgagctgt 960  
|||||  
Db 901 aagagcacaataccgagctgt 960  
QY 961 gctgtgacacatgtcgagagagcagcttcgacgagcagcagcagcagcagcagcagcagcagc 1020  
|||||  
Db 961 gctgtgacacatgtcgagagagcagcttcgacgagcagcagcagcagcagcagcagcagcagc 1020  
QY 1021 gcaagcccaatcgagacatctgagagatcagcagcagcagcagcagcagcagcagcagcagc 1080  
|||||  
Db 1021 gcaagcccaatcgagacatctgagagatcagcagcagcagcagcagcagcagcagcagcagc 1080  
QY 1081 ggcgactcactgt 1140  
|||||  
Db 1081 ggcgactcactgt 1140  
QY 1141 atcatgtctcggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200  
|||||  
Db 1141 atcatgtctcggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200  
QY 1201 ccacagctatgagcagcttcataggt 1260  
|||||  
Db 1201 ccacagctatgagcagcttcataggt 1260  
QY 1261 ggaactcagcagcagctgt 1320  
|||||  
Db 1261 ggaactcagcagcagctgt 1320  
QY 1321 cgaagagctgagcagcttcataggt 1380  
|||||  
Db 1321 cgaagagctgagcagcttcataggt 1380  
QY 1381 catctccagcagcagcttcataggt 1440  
|||||  
Db 1381 catctccagcagcagcttcataggt 1440  
QY 1441 tgaagccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1500  
|||||  
Db 1441 tgaagccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1500  
QY 1501 ggggaaaaagcagcagctgt 1560  
|||||  
Db 1501 ggggaaaaagcagcagctgt 1560  
QY 1561 gcaacacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1620  
|||||

Db 1561 gcaacacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1620  
QY 1621 ggt 1680  
|||||  
Db 1621 ggt 1680  
QY 1681 caacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1740  
|||||  
Db 1681 caacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1740  
QY 1741 acaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1800  
|||||  
Db 1741 acaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1800  
QY 1801 ccacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1860  
|||||  
Db 1801 ccacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1860  
QY 1861 cccggagatgagtaataagatagagagagcagcagcagcagcagcagcagcagcagcagcagc 1894  
|||||  
Db 1861 cccggagatgagtaataagatagagagagcagcagcagcagcagcagcagcagcagcagcagc 1894

RESULT 8  
US-09-064-645-14  
Sequence 14, Application US/09064645  
GENERAL INFORMATION:  
APPLICANT: PATRICIA A. BILLING-MEDEL  
APPLICANT: MAURICE COHEN  
APPLICANT: TRACEY L. COLPITTS  
APPLICANT: PAULA N. FRIEDMAN  
APPLICANT: MICHAEL R. KLASS  
APPLICANT: LISA ROBERTS-RAPP  
APPLICANT: JOHN C. RUSSELL  
APPLICANT: STEPHEN D. STROUPE  
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL  
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/064, 645  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/845, 995  
FILING DATE: 25-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Becker, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 6085, US, .p1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847/935-1729  
TELEFAX: 847/938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1902 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-064-645-14









|||||  
Db 1268 ggaactagcccccgggtggcccgctctgtggtgacatccgcaagaacccgtcccgcatbaacta 1327  
Qy 1321 cgaacaagctgagccgctccatccatccgacagctctacaagaaggagcatccatccggaaccaga 1380  
Db 1328 cgaacaagctgagccgctccatccatccgacagctctacaagaaggagcatccatccggaaccaga 1387  
Qy 1381 catctccagcgcctctgtctacacagttctgtacaccccatctgagtgctgcccagggcc 1440  
Db 1388 catctccagcgcctctgtctacacagttctgtacaccccatctgagtgctgcccagggcc 1447  
Qy 1441 tgaaccccgccctcagggccctctctctgctgacctgcccacgacgaagccctagatg 1500  
Db 1448 tgaaccccgccctcagggccctctctctgctgacctgcccacgacgaagccctagatg 1507  
Qy 1501 ggggaaaaagcggcagtgctgctctgctgctgacctcctcagagcccaaggtcagggaggg 1560  
Db 1508 ggggaaaaagcggcagtgctgctctgctgctgacctcctcagagcccaaggtcagggaggg 1567  
Qy 1561 gcaaccaactgcccccgaggagataatggtccctctgagccctcgggacataagggaggg 1620  
Db 1568 gcaaccaactgcccccgaggagataatggtccctctgagccctcgggacataagggaggg 1627  
Qy 1621 gtccttcctcctcagggcccgagctgctcccttgaggagagaggagagagagctgctccc 1680  
Db 1628 gtccttcctcctcagggcccgagctgctcccttgaggagagaggagagagagctgctccc 1687  
Qy 1681 caaacactgctctgacccccagcatctccagagcagagcctacagaaggcagtgactcg 1740  
Db 1688 caaacactgctctgacccccagcatctccagagcagagcctacagaaggcagtgactcg 1747  
Qy 1741 acaaaagccacagcagctccagcctctctctgctccatcccccgtgctcccatctgca 1800  
Db 1748 acaaaagccacagcagctccagcctctctctgctccatcccccgtgctcccatctgca 1807  
Qy 1801 ccacacctgagtggtgcagaggagacatctgcacccctagttggcagcagggagtgcc 1860  
Db 1808 ccacacctgagtggtgcagaggagacatctgcacccctagttggcagcagggagtgcc 1867  
Qy 1861 cccgggaatgagataaagatactagagaactg 1894  
Db 1868 cccgggaatgagataaagatactagagaactg 1901

RESULT 11  
US-09-567-970-15  
; Sequence 15, Application US/09567970  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Billing-Medel, Patricia A.  
; APPLICANT: Colpitts, Tracey L.  
; APPLICANT: Friedman, Paula N.  
; APPLICANT: Klass, Michael R.  
; APPLICANT: Roberts-Rapp, Lisa  
; APPLICANT: Russell, John C.  
; APPLICANT: Stroupe, Stephen D.  
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR  
; FILE REFERENCE: 6085, US, P2  
; CURRENT APPLICATION NUMBER: US/09/567, 970  
; CURRENT FILING DATE: 2000-05-10  
; PRIOR APPLICATION NUMBER: US 09/064, 645  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: US 08/845, 995  
; PRIOR FILING DATE: 1997-04-25  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 1902  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-567-970-15

Query Match 98.6%; Score 1868.4; DB 22; Length 1902;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1878; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
Qy 1 gtcgtacctctccagagacatctctgacatctccgtgtctacacatgccccacagagccc 60  
Db 8 gtcgtacctctccagagacatctctgacatctccgtgtctacacatgccccacagagccc 67  
Qy 61 agctctccaaagcctgctctccagctccctcctcagagccctcagagttgggctctgacaggtg 120  
Db 68 agctctccaaagcctgctctccagctccctcctcagagccctcagagttgggctctgacaggtg 127  
Qy 121 ccagagcagcagccctgagctggggtlaaggagatccctacagagcagcagccctgagagc 180  
Db 128 ccagagcagcagccctgagctggggtlaaggagatccctacagagcagcagccctgagagc 187  
Qy 181 gccaaccaacccctctgagagtggtgccaagggccccaagtggccaacccctgagagctctc 240  
Db 188 tcaagaggccacccctctgagagtggtgccaagggccccaagtggccaacccctgagagctctc 247  
Qy 241 gccaaccaacccctgagagtggtgccaagggccccaagtggccaacccctgagagcagc 300  
Db 248 gccaaccaacccctgagagtggtgccaagggccccaagtggccaacccctgagagcagc 307  
Qy 301 cagtgagctcagctgagccacacctctcccgagccctgagagtggtgagctgacagcagagc 360  
Db 308 cagtgagctcagctgagccacacctctcccgagccctgagagtggtgagctgacagcagagc 367  
Qy 361 ctcccttgagcagcagcagcttaacagagacagcccgagcagcccaacagcagcagcagc 420  
Db 368 ctcccttgagcagcagcagcttaacagagacagcccgagcagcccaacagcagcagcagc 427  
Qy 421 cagcgcagcagcccggtctgagcagcagctatccccaagcacaactctgctgcccccgagc 480  
Db 428 cagcgcagcagcccggtctgagcagcagctatccccaagcacaactctgctgcccccgagc 487  
Qy 481 ggtgtcgcagcagcagcttgagagagcgagcagcggtggcagtggtgtctcagagagacgga 540  
Db 488 ggtgtcgcagcagcagcttgagagagcgagcagcggtggcagtggtgtctcagagagacgga 547  
Qy 541 ctggaggtccagctccaccccgagcagcccgagcagagcctgtccgctctctacacccctcta 600  
Db 548 ctggaggtccagctccaccccgagcagcccgagcagagcctgtccgctctctacacccctcta 607  
Qy 601 ctgtgacatgctgtaaccttgagagcagcagctgggacagcagcccttggtggcagcag 660  
Db 608 ctgtgacatgctgtaaccttgagagcagcagctgggacagcagcccttggtggcagcag 667  
Qy 661 tcggagagagcacaactgacagcagctgagcagtgcccggtcatctgacagcacaagcccgagc 720  
Db 668 tcggagagagcacaactgagcagcagctgagcagtgcccggtcatctgacagcacaagcccgagc 727  
Qy 721 gggcaccctgagacttggtgcccggcgagctgaccttgagagagacactccgctggagcaggt 780  
Db 728 gggcaccctgagacttggtgcccggcgagctgaccttgagagagacactccgctggagcaggt 787  
Qy 781 gcagtccatggtggtgagcagagtgctcaagagacatcagagcagcctgcaagcgtccaa 840  
Db 788 gcagtccatggtggtgagcagagtgctcaagagacatcagagcagcctgcaagcgtccaa 847  
Qy 841 catcaacgacatcccatgagctgagagcccgagcaaatgtgcagaagtgtgctcctgtgac 900  
Db 848 catcaacgacatcccatgagctgagagcccgagcaaatgtgcagaagtgtgctcctgtgac 907  
Qy 901 agagacccaataacggctgccccccatgaggagagcccttcagagagctgggggcaagga 960  
Db 908 agagacccaataacggctgccccccatgaggagagcccttcagagagctgggggcaagga 967  
Qy 961 gctgtgcacatgtctcagagagcagttccgcagcagcgtcgcagccctgggtggagtgct 1020  
Db 968 gctgtgcacatgtctcagagagcagttccgcagcagcgtcgcagccctgggtggagtgct 1027

QY	1021	gcaagcccaactggaacacatcugaaatcgaatcgagccgctggaatgaagaagcgagacttaactctg	1080
Db	1028	gcaagcccaactctggaacactctggaatgaatcgaatcgagccctctggaatgaagaagcgagacttaactctg	1087
QY	1081	ggcgatccaactgctgctccctcgagaccagctgagagagactctgacccgaacagccagcgctgagctc	1140
Db	1088	ggcgatccaactgctgctccctcgagaccagctgagagagactctgacccgaacagccagcgagctgagctc	1147
QY	1141	atcatgctccgagcgagcccaatccacctgtgcaatctctccctcaagagaggttgcctactcaagcc	1200
Db	1148	atcatgctccgagcgagcccaatccacctgtgcaatctctccctcaagagaggttgcctactcaagcc	1207
QY	1201	cccaagactaibgcccgtctcaatibgtgtgctccaacaaagagaaagggcatcttccaatatga	1260
Db	1208	cccaagactaibgcccgtctcaatibgtgtgctccaacaaagagaaagggcatcttccaatatga	1267
QY	1261	ggactccaagcccaaggttgcccccgtctgtggtgggcatcccgcaagaaacgctcccgcaatgaatg	1320
Db	1268	ggactccaagcccaaggttgcccccgtctgtggtgggcatcccgcaagaaacgctcccgcaatgaatg	1327
QY	1321	cgacaagactgagcccgccccaatcccgccaagctcttaacaagaagggcatcatcccggaagcaga	1380
Db	1328	cgacaagactgagcccgccccaatcccgccaagctcttaacaagaagggcatcatcccggaagcaga	1387
QY	1381	catctccaagcgcctgctgctacaagattcgbgacccccatctgaagtgccttggcccaagggcc	1440
Db	1388	catctccaagcgcctgctgctacaagattcgbgacccccatctgaagtgccttggcccaagggcc	1447
QY	1441	tggaaaacggccctcaaggggacctctctccctcgccctcgccctcgaccccaagccagcccgagatg	1500
Db	1448	tggaaaacggccctcaaggggacctctctccctcgccctcgccctcgaccccaagccagcccgagatg	1507
QY	1501	gggggaaaaacgggcaagtgtgctctgctgcgtctctgaaccttcccaagggcccaaggtccaagggaggg	1560
Db	1508	gggggaaaaacgggcaagtgtgctctgctgcgtctctgaaccttcccaagggcccaaggtccaagggaggg	1567
QY	1551	gcaaaccaactgcgcccagggagggatataibgtgtctctctgagggcccttcggagccaatibgagaggg	1620
Db	1558	gcaaaccaactgcgcccagggagggatataibgtgtctctctgagggcccttcggagccaatibgagaggg	1627
QY	1621	gtgctctccctccaagggccagcgtgcctccctctggaagaagaagggagagagcgctgctccc	1680
Db	1628	gtgctctccctccaagggccagcgtgcctccctctggaagaagaagggagagagcgctgctccc	1687
QY	1681	caaacactgctctctgaaccccaagcatcttccgaagagcgactcaaaagagggcgagtactcg	1740
Db	1688	caaacactgctctctgaaccccaagcatcttccgaagagcgactcaaaagagggcgagtactcg	1747
QY	1741	acaaagggccacagggcagttccaagcgctctctgtgctcaatccacctgcctcccatctctga	1800
Db	1748	acaaagggccacagggcagttccaagcgctctctgtgctcaatccacctgcctcccatctctga	1807
QY	1801	cgacaacctggaatgctgcagaaggagaaatctgcacccctcaagtctggcgagcgagaggtgcc	1860
Db	1808	cgacaacctggaatgctgcagaaggagaaatctgcacccctcaagtctggcgagcgagaggtgcc	1867
QY	1861	cccggaatgatatataaataactagagaactg	1894
Db	1868	cccggaatgatatataaataactagagaactg	1901
RESULT 12			
US-09-718-159-1			
; Sequence 1, Application US/09718159			
; GENERAL INFORMATION:			
; APPLICANT: Harkins, Richard			
; APPLICANT: Lin, Richard			
; APPLICANT: Luke, May			
; APPLICANT: Montecarlo, Felipe			
; APPLICANT: Parkes, Deborah			
; APPLICANT: Parry, Gordon			
; APPLICANT: Steindreher, Renate			
; APPLICANT: Van Heult, Pamela Toy			

```

1  APPLICANT: Xuan, Jian-ai
2  TITLE OF INVENTION: DNA Encoding a Novel ETS Polypeptide
3  FILE REFERENCE: 51790AUSM1
4  CURRENT APPLICATION NUMBER: US/09/718,159
5  CURRENT FILING DATE: 2000-11-21
6  PRIOR APPLICATION NUMBER: 60/166,182
7  PRIOR FILING DATE: 1999-11-30
8  NUMBER OF SEQ ID NOS: 11
9  SOFTWARE: PatentIn Ver. 2.0
10 SEQ ID NO 1
11 LENGTH: 1918
12 TYPE: DNA
13 ORGANISM: Homo sapiens
14 FEATURE:
15 NAME/KEY: CDS
16 LOCATION: (422)..(1429)
17 US-09-718-159-1
18
19 Query Match          98.6%; Score 1868.4; DB 28; Length 1918;
20 Best Local Similarity 99.2%; Pred. No. 0;
21 Matches 1878; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
22
23 QY 1 gctctgactctctccagcaacattccttgacatctgcccgtgtccacaactgcccacagacc 60
24 Db 7 gctctgactctctccagcaacattccttgacatctgcccgtgtccacaactgcccacagacc 66
25
26 QY 61 atctctcccaagcctgtctgcaagctctctctccaagccctctagtttggtcttgccagctgt 120
27 Db 67 agtctctccaagcctgtctgcaagctctctctccaagccctctagtttggtcttgccagctgt 126
28
29 QY 121 ccagcagcagcagccctgtggtgtctgggtgtaggtgagctccctacaaagcagcagccctgagacc 180
30 Db 127 ccagcagcagcagccctgtggtgtctgggtgtaggtgagctccctacaaagcagcagccctgagacc 186
31
32 QY 181 gccacacagccacccctctgaggggtgtgcagggcccccagttgccaacctgtgactgtcctct 240
33 Db 187 tcagagagcagccacccctctgaggggtgtgcagggcccccagttgccaacctgtgactgtcctct 246
34
35 QY 241 gccacacagccctctgtgcccctgtgttcgcgctgtgccccccagatgtcttgctgtagaacagc 300
36 Db 247 gccacacagccctctgtgcccctgtgttcgcgctgtgccccccagatgtcttgctgtagaacagc 306
37
38 QY 301 cagttgacctcagctgtgcaccaactctctccggcgccctctgtaggtgtgtgacatgcagcagacag 360
39 Db 307 cagttgacctcagctgtgcaccaactctctccggcgccctctgtaggtgtgtgacatgcagcagacag 366
40
41 QY 361 ctccctgtggcaccaggtcagctaacagacacagcgcgcacagcccaaacagcagcgcgtatgg 420
42 Db 367 ctccctgtggcaccaggtcagctaacagacacagcgcgcacagcccaaacagcagcgcgtatgg 426
43
44 QY 421 cagcgccagcccccgggtgtctgtagcagcgtatactccccaagcacaactccgtgtctccccccagacc 480
45 Db 427 cagcgccagcccccgggtgtctgtagcagcgtatactccccaagcacaactccgtgtctccccccagacc 486
46
47 QY 481 ggtgtctgcggaacagcctltgtagaagcagcgcagccggtgggtcagctgtgtctctgtagaagacgga 540
48 Db 487 ggtgtctgcggaacagcctltgtagaagcagcgcagccggtgggtcagctgtgtctctgtagaagacgga 546
49
50 QY 541 ctgtagagtcaccagttccaaccgcgcagcccgagcaggtgctgtccgctcttgacactctctcta 600
51 Db 547 ctgtagagtcaccagttccaaccgcgcagcccgagcaggtgctgtccgctcttgacactctctcta 606
52
53 QY 601 ctttgacaatgctgtatacctctgtagaagacagcagcgtcgggtgagcagaagggccctctgggtccagcag 660
54 Db 607 ctttgacaatgctgtatacctctgtagaagacagcagcgtcgggtgagcagaagggccctctgggtccagcag 666
55
56 QY 661 tcgggagagagccacacctgagaccagctctgtagagcgtgcccgtgtcatcttgacaagcccaagcccaagc 720
57 Db 667 tcgggagagagccacacctgagagagcctctgtagagcgtgcccgtgtcatcttgacaagcccaagcccaagc 726
58
59 QY 721 gggacacccctgtagctgtgtgccccggcggtgtgacctltgtagaggaacactgctctgtagaagcagt 780
60 Db 727 gggacacccctgtagctgtgtgccccggcggtgtgacctltgtagaggaacactgctctgtagaagcagt 786

```

[illegible]

```

QY 1861 cccgggaatgataataagaatactcagaagactg 1894
      |||||||
Db 1867 cccgggaatgataataagaatactcagaagactg 1900

RESULT 13
US-09-570-593-12
; Sequence 12, Application US/09570593
; GENERAL INFORMATION:
; APPLICANT: Kaufmann, Joerg
; APPLICANT: Xin, Hong
; APPLICANT: Harrowe, Greg
; TITLE OF INVENTION: EXPRESSION OF ETS-DOMAIN PROTEINS IN
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 2300-1556
; CURRENT APPLICATION NUMBER: US/09/570,593
; CURRENT FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 60/134,112
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 3317
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)...(3317)
; OTHER INFORMATION: GSEF-encoding sequence with promoter (Figs. 2A-2B)
; NAME/KEY: promoter
; LOCATION: (1)...(1380)
; OTHER INFORMATION: Promoter
; NAME/KEY: CDS
; LOCATION: (1815)...(2819)
; OTHER INFORMATION: GSEF coding sequence
US-09-570-593-12

Query Match 98.5%, Score 1865.2; DB 22; Length 3317;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1876; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 gctgactctccccaagaatctctgtcaacttgcggtgtccaacttgcaccaagacc 60
      |||||||
Db 1400 gctgactctccccaagaatctctgtcaacttgcggtgtccaacttgcaccaagacc 1459

QY 61 agtctccaagctctgtccaaagtcccttccaagccctcagaatttgagcttgcaagtg 120
      |||||||
Db 1460 agtctccaagctctgtccaaagtcccttccaagccctcagaatttgagcttgcaagtg 1519

QY 121 ccagcagcagcctctgggctgggtaagggaactccctacagcagcagccctgagacc 180
      |||||||
Db 1520 ccagcagcagcctctgggctgggtaagggaactccctacagcagcagccctgagacc 1579

QY 181 gccaccagcaccctctgaggttgccagggcccccagtgccaacttgaagtctcctct 240
      |||||||
Db 1580 tcaagagggcaccctctgaggttgccagggcccccagtgccaacttgaagtctcctct 1639

QY 241 gccaccagcctctgctggccctctgttcacgcttgcccccagatgcttgctagaacac 300
      |||||||
Db 1640 gccaccagcctctgctggccctctgttcacgcttgcccccagatgcttgctagaacac 1699

QY 301 cagtgagctcagctgcccaacctctcccgcccttgagttggaacttgacagacaag 360
      |||||||
Db 1700 cagtgagctcagctgcccaacctctcccgcccttgagttggaacttgacagacaag 1759

QY 361 ctcccttgagcaccagcagctacacagaaacacgcgcagcccaaacagcagcgatggg 420
      |||||||
Db 1760 ctcccttgagcaccagcagctacacagaaacacgcgcagcccaaacagcagcgatggg 1819

QY 421 cagcgccaacccgggtctttagcagcgtatcccccagccacccctcctgctccccccaac 480

```

Dh 1820 cagcgccagcccggtctgtagcagctatccccagccactctgctccccccgaac 1879  
Qy 481 ggtctcgagcaagcctctgagaaagcagcagcgaggcagctggtctcgagaaacgga 540  
Dh 1880 ggtctcgagcaagcctctgagaaagcagcagcgaggcagctggtctcgagaaacgga 1339  
Qy 541 ctgtagtccagctccacccgcaacccgagcagagcctctcgagcttcgaacttcta 600  
Dh 1940 ctgtagtccagctccacccgcaacccgagcagagcctctcgagcttcgaacttcta 1999  
Qy 601 ctctgagactctgtagcagcagcagcagcagcagcagcagcagcagcagcagcag 660  
Dh 2000 ctctgagactctgtagcagcagcagcagcagcagcagcagcagcagcagcagcag 2059  
Qy 661 tccggagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
Dh 2060 tccggagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2119  
Qy 721 gggcagcctgagctggtgctgctgctgctgctgctgctgctgctgctgctgctgct 780  
Dh 2120 gggcagcctgagctggtgctgctgctgctgctgctgctgctgctgctgctgctgct 2179  
Qy 781 ggcagtcacgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 840  
Dh 2180 ggcagtcacgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 2239  
Qy 841 catcacgcagatcccatgagactgtagagcccaagcaatgtagaagtggtctctgtgac 900  
Dh 2240 catcacgcagatcccatgagactgtagagcccaagcaatgtagaagtggtctctgtgac 2299  
Qy 901 agagcacaacaatacagctgctgcccccatggtggaagccttcagagagctgctgagga 960  
Dh 2300 agagcacaacaatacagctgctgcccccatggtggaagccttcagagagctgctgagga 2359  
Qy 961 gctctgagcagctgctgagagagcagcttccgcaagcagctcgcctggtggtggtggt 1020  
Dh 2360 gctctgagcagctgctgagagagcagcttccgcaagcagctcgcctggtggtggtggt 2419  
Qy 1021 gcaagcccaactgtagactgtagaagtcagcagcctgtagaagagcagcagcagcag 1080  
Dh 2420 gcaagcccaactgtagactgtagaagtcagcagcctgtagaagagcagcagcagcag 2479  
Qy 1081 ggcagatcacactggtgctctgtagcagctgtagagagagctgtagcagcagcagcag 1140  
Dh 2480 ggcagatcacactggtgctctgtagcagctgtagagagagctgtagcagcagcagcag 2539  
Qy 1141 atcatgtctcgagcagcagcagcagcagctgtagcagcttccctcaagagctgtagcagc 1200  
Dh 2540 atcatgtctcgagcagcagcagcagcagctgtagcagcttccctcaagagctgtagcagc 2599  
Qy 1201 ccacagctatgagcagcttcatlagtggtcctaacaagagagagcagcttctcaaatgta 1260  
Dh 2600 ccacagctatgagcagcttcatlagtggtcctaacaagagagagcagcttctcaaatgta 2659  
Qy 1261 ggcagtcagccagagtggtgctgctggtggtggtggtggtggtggtggtggtggtggt 1320  
Dh 2660 ggcagtcagccagagtggtgctgctggtggtggtggtggtggtggtggtggtggtggt 2719  
Qy 1321 cgaacaagctgagcagcttccatcagcagcttctcaacaagagagcagcttccggaagcaga 1380  
Dh 2720 cgaacaagctgagcagcttccatcagcagcttctcaacaagagagcagcttccggaagcaga 2779  
Qy 1381 catctccagcagctggtctatcagagcttctgagccacatctgagctggtggtggtggt 1440  
Dh 2780 catctccagcagctggtctatcagagcttctgagccacatctgagctggtggtggtggt 2839  
Qy 1441 tgaacccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1500  
Dh 2840 tgaacccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2899  
Qy 1501 gggagaaaagcagcagctggtctgctgctgctgctgctgctgctgctgctgctgctgctg 1560  
Dh 2900 gggagaaaagcagcagctggtctgctgctgctgctgctgctgctgctgctgctgctgctg 2959

Qy 1561 gcaaccaactgcccccaaggagatagtggtctctctggtggtcctcgagacatgaggcagag 1620  
Dh 2960 gcaaccaactgcccccaaggagatagtggtctctctggtggtcctcgagacatgaggcagag 3019  
Qy 1621 gttgtctctctcagcagcagctggtctctctctgtagagagagagagagagagctgctcc 1680  
Dh 3020 gttgtctctctcagcagcagctggtctctctctgtagagagagagagagagagctgctcc 3079  
Qy 1681 caacacctctctctgaccccaagcatttccagagcagagcctcagaagagagagagagag 1740  
Dh 3080 caacacctctctctgaccccaagcatttccagagcagagcctcagaagagagagagagag 3139  
Qy 1741 acaaaagccacagcagcagctcctctctctctctctctctctctctctctctctctctca 1800  
Dh 3140 acaaaagccacagcagcagctcctctctctctctctctctctctctctctctctctctca 3199  
Qy 1801 ccacacctggtatggtgtagagagatctgcaacctgaagtgtggtggtggtggtggtggt 1860  
Dh 3200 ccacacctggtatggtgtagagagatctgcaacctgaagtgtggtggtggtggtggtggt 3259  
Qy 1861 cccgggaatggataataagatactagagagctg 1894  
Dh 3260 cccgggaatggataataagatactagagagctg 3293

RESULT 14  
US-09-866-356-2  
; Sequence 2, Application US/09866356  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
Corley, Neil C.  
Guegler, Karl J.  
Lal, Preeti  
TITLE OF INVENTION: PROSTATE-ASSOCIATED ETS PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: PastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/866,356  
FILING DATE: 29-May-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/055,113  
FILING DATE: 1998-04-03  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0501 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1905 base pairs  
TYPE: nucleic acid  
STRANDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PROSTOT12  
CLONE: 1813005  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-866-356-2





SOFTWARE: PERL Program  
; SEQ ID NO 74  
; LENGTH: 1907  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 234340.7  
US-60-209-009-74

Query Match 97.0%; Score 1837.8; DB 59; Length 1907;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1870; Conservative 0; Mismatches 17; Indels 2; Gaps 2;

QY 1 gcttgatctctcccaagcaatctctgcaactctgcgcgtgtccaaactgtcccaagacc 60  
Db 15 gtctgactctctcccaagcaatctctgcaactctgcgcgtgtccaaactgtcccaagacc 74  
QY 61 agctctccaaagctgtgccaagct-cctctgcaagccctccaaagtgtgacctgccaagt 119  
Db 75 agctctccaaagctgtgccaagctgacctgcaagccctccaaagtgtgacctgccaagt 134  
QY 120 gccacgaagcagcctt-gggctggggtgtagggactccctacagcagcagcctctgaga 178  
Db 135 gccacgaagcagccttggggtgttagggactccctacagcagcagcctctgaga 194  
QY 179 ccgcaccac 238  
Db 195 ccgcaccac 254  
QY 239 ctgcac 298  
Db 255 ctgcac 314  
QY 299 gccacgtggacctgacctgacctgacctgacctgacctgacctgacctgacctgacctg 358  
Db 315 gccacgtggacctgacctgacctgacctgacctgacctgacctgacctgacctgacctg 374  
QY 359 agctctctgggac 418  
Db 375 agctctctgggac 434  
QY 419 ggcacgac 478  
Db 435 ggcacgac 494  
QY 479 agcgtgtcgggac 538  
Db 495 agcgtgtcgggac 554  
QY 539 gactggagctcccaagctcccaagctcccaagctcccaagctcccaagctcccaagctcc 598  
Db 555 gactggagctcccaagctcccaagctcccaagctcccaagctcccaagctcccaagctcc 614  
QY 599 tactttgacatgtgtac 658  
Db 615 tactttgacatgtgtac 674  
QY 659 agctcggagagac 718  
Db 675 agctcggagagac 734  
QY 719 gccgggac 778  
Db 735 gccgggac 794  
QY 779 gtgcagac 838  
Db 795 gtgcagac 854  
QY 839 aacatcac 898

Db 855 aacatcac 914  
QY 899 aacaaagac 958  
Db 915 aacaaagac 974  
QY 959 gactgtgac 1018  
Db 975 gactgtgac 1034  
QY 1019 ctgcacgac 1078  
Db 1035 ctgcacgac 1094  
QY 1079 gggcgatctac 1138  
Db 1095 gggcgatctac 1154  
QY 1139 tcatcatgtctcgggac 1198  
Db 1155 tcatcatgtctcgggac 1214  
QY 1199 ccccaac 1258  
Db 1215 ccccaac 1274  
QY 1259 gagagac 1318  
Db 1275 gagagac 1334  
QY 1319 tgcgaac 1378  
Db 1335 tgcgaac 1394  
QY 1379 gacatctccac 1438  
Db 1395 gacatctccac 1454  
QY 1439 ccttgaaac 1498  
Db 1455 ccttgaaac 1514  
QY 1499 tgggagaaac 1558  
Db 1515 tgggagaaac 1574  
QY 1559 gggcacaac 1618  
Db 1575 gggcacaac 1634  
QY 1619 ggggtgtctctcctcagac 1678  
Db 1635 ggggtgtctctcctcagac 1694  
QY 1679 cccaac 1738  
Db 1695 cccaac 1754  
QY 1739 cgacaagac 1798  
Db 1755 cgacaagac 1814  
QY 1799 caccac 1858  
Db 1815 caccac 1874  
QY 1859 ccccgaggac 1887  
Db 1875 ccccgaggac 1903



Wed Aug 7 09:11:23 2002

us-09-841-963a-1.rnp

Page 19

Job time: 9120 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 19:53:21 ; Search time 3498.6 Seconds  
(without alignments)  
11328.777 Million cell updates/sec

Title: US-09-841-963a-1  
1894  
Sequence: 1 gctgactcctccagcac.....ataagatactagagaactg 1894

Scoring table: IDENTITY-NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 705160

Minimum DB seq length: 10  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBank:  
1: gb\_ba:\*  
2: gb\_hvg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_mu:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_hvg\_hum:\*  
31: em\_hvg\_inv:\*  
32: em\_hvg\_other:\*  
33: em\_hlgo\_inv:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

C	1	34	1.8	34	6	AX048158	AX048158 Sequence
	2	31	1.6	31	6	AX048157	AX048157 Sequence
	3	23	1.2	23	6	AX175520	AX175520 Sequence
	4	23	1.2	23	6	AX175524	AX175524 Sequence
	5	23	1.2	23	6	AX175525	AX175525 Sequence
	6	23	1.2	23	6	AX175526	AX175526 Sequence
	7	22.8	1.2	50	6	AX199688	AX199688 Sequence
	8	22.6	1.2	48	6	AX221945	AX221945 Sequence
	9	21.6	1.1	31	6	AX249190	AX249190 Sequence
C	10	21.4	1.1	49	10	MUSJHCD6	M64562 Mouse JH-cd
	11	21.2	1.1	45	6	A05116	A05116 Oligonucleo
	12	21.2	1.1	48	6	HSTCK6x45	269575 H.sapiens m
	13	21.2	1.1	48	9	HSTCK6x45	269547 H.sapiens m
C	14	21.2	1.1	50	6	AX250311	AX250311 Sequence
	15	21.2	1.1	21	6	AX175518	AX175518 Sequence
C	16	21	1.1	37	6	AX139078	AX139078 Sequence
	17	21	1.1	41	6	AX327072	AX327072 Sequence
	18	20.8	1.1	48	6	AX221416	AX221416 Sequence
	19	20.8	1.1	48	6	AX221427	AX221427 Sequence
	20	20.8	1.1	48	6	AX274492	AX274492 Sequence
	21	20.8	1.1	49	6	AX279693	AX279693 Sequence
C	22	20.8	1.1	50	6	AR123173	AR123173 Sequence
C	23	20.8	1.1	50	6	AX165853	AX165853 Sequence
	24	20.6	1.1	48	6	AX221428	AX221428 Sequence
C	25	20.6	1.1	48	6	AX320747	AX320747 Sequence
	26	20.6	1.1	50	9	HUMTCRVJ46	AX320747 Sequence
	27	20.6	1.1	48	6	AX221643	L39584 Homo sapien
	28	20.4	1.1	48	6	AX221644	AX221644 Sequence
	29	20.4	1.1	48	6	AX221665	AX221665 Sequence
	30	20.4	1.1	48	6	AX274490	AX274490 Sequence
	31	20.4	1.1	48	6	AX274552	AX274552 Sequence
	32	20.4	1.1	50	6	AR003562	AR003562 Sequence
	33	20.4	1.1	50	6	AX057285	AX057285 Sequence
C	34	20.2	1.1	37	6	A46812	A46812 Sequence 9
	35	20.2	1.1	42	6	AR021417	AR021417 Sequence
	36	20.2	1.1	42	6	AR042979	AR042979 Sequence
	37	20.2	1.1	42	6	AR161282	AR161282 Sequence
	38	20.2	1.1	42	6	I43661	I43661 Sequence 50
	39	20.2	1.1	42	6	I62970	I62970 Sequence 67
	40	20.2	1.1	42	6	I88723	I88723 Sequence 67
	41	20.2	1.1	45	9	S63845	S63845 (DJ/TM4-del
C	42	20.2	1.1	46	6	AR140182	AR140182 Sequence
C	43	20.2	1.1	46	6	AR173249	AR173249 Sequence
C	44	20	1.1	36	6	AX007193	AX007193 Sequence
C	45	20	1.1	36	6	AX007365	AX007365 Sequence

## ALIGNMENTS

RESULT 1  
AX048158/c  
LOCUS AX048158 34 bp DNA  
DEFINITION Sequence 11 from Patent WO0070092.  
ACCESSION AX048158  
VERSION AX048158.1 GI:11876974  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
artificial sequence.  
REFERENCE  
1 (bases 1 to 34)  
AUTHORS Kaufmann, J. and Xin, H.  
TITLE Expression of ets-domain proteins in cancer  
JOURNAL Patent: WO 0070092-A 11 23-NOV-2000;  
Chiron Corporation (US)  
FEATURES  
source  
1..34  
Location/Qualifiers  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="GSEF-specific primer"

BASE COUNT	5	a	10	c	7	g	12	t
------------	---	---	----	---	---	---	----	---

Query Match 1.8%; Score 34; DB 6; Length 34;  
Best Local Similarity 100.0%; Pred. No. 7.1e+04;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1850 ccagagatgccccgggaatgataaagata 1883  
|||||  
DB 34 CCAGAGATGCCCCGGGAATGATATAAGATA 1

RESULT 2  
AX048157 AX048157 31 bp DNA linear PAT 15-DEC-2000  
LOCUS Sequence 10 from Patent WO0070092.  
DEFINITION AX048157  
ACCESSION AX048157  
VERSION AX048157.1 GI:11876973  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 31)  
AUTHORS Kaufmann, J. and Xin, H.  
TITLE Expression of ets-domain proteins in cancer  
JOURNAL Patent: WO 0070092-A 10 23-NOV-2000;  
Chiron Corporation (US)  
FEATURES  
source Location/Qualifiers  
1..31  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="GSEF-specific primer"

BASE COUNT 8 a 9 c 13 g 1 t  
ORIGIN  
Query Match 1.6%; Score 31; DB 6; Length 31;  
Best Local Similarity 100.0%; Pred. No. 2.3e+05;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1552 caggaagggcaaccactgccccagggga 1582  
|||||  
DB 1 CAGGAGCGGCGACACACTGCCCGAGCGGA 31

RESULT 3  
AX175520 AX175520 23 bp DNA linear PAT 03-JUL-2001  
LOCUS Sequence 5 from Patent WO0142472.  
DEFINITION AX175520  
ACCESSION AX175520  
VERSION AX175520.1 GI:14598659  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 23)  
AUTHORS Harkins, R., Lin, R.J., Luke, M., Monteciaro, F., Parkes, D., Parry, G.,  
Steindbrecher, R., van Heult, P.T. and Xuan, J.A.  
TITLE Dna encoding a novel prost-ets polypeptide  
JOURNAL Patent: WO 0142472-A 5 14-JUN-2001;  
SCHERING AKTIENGESellschaft (DE)  
FEATURES  
source Location/Qualifiers  
1..23  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="probe"

BASE COUNT 4 a 12 c 2 g 5 t  
ORIGIN

Query Match 1.2%; Score 23; DB 6; Length 23;  
Best Local Similarity 100.0%; Pred. No. 5.4e+06;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1783 cctgctcccatctgtgaccaca 1805

DB 1 CCTGCTCCCATCTGTGACCACA 23  
|||||

RESULT 4  
AX175524 AX175524 23 bp mRNA linear PAT 03-JUL-2001  
LOCUS Sequence 9 from Patent WO0142472.  
DEFINITION AX175524  
ACCESSION AX175524  
VERSION AX175524.1 GI:14598660  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 23)  
AUTHORS Harkins, R., Lin, R.J., Luke, M., Monteciaro, F., Parkes, D., Parry, G.,  
Steindbrecher, R., van Heult, P.T. and Xuan, J.A.  
TITLE Dna encoding a novel prost-ets polypeptide  
JOURNAL Patent: WO 0142472-A 9 14-JUN-2001;  
SCHERING AKTIENGESellschaft (DE)  
FEATURES  
source Location/Qualifiers  
1..23  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="antisense"

BASE COUNT 4 a 8 c 5 g 6 t  
ORIGIN

Query Match 1.2%; Score 23; DB 6; Length 23;  
Best Local Similarity 100.0%; Pred. No. 5.4e+06;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1161 tccacctgtgagctcctcaag 1183  
|||||  
DB 1 TCCACCTGTGGCAGTCTCTCAAG 23

RESULT 5  
AX175525 AX175525 23 bp mRNA linear PAT 03-JUL-2001  
LOCUS Sequence 10 from Patent WO0142472.  
DEFINITION AX175525  
ACCESSION AX175525  
VERSION AX175525.1 GI:14598661  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 23)  
AUTHORS Harkins, R., Lin, R.J., Luke, M., Monteciaro, F., Parkes, D., Parry, G.,  
Steindbrecher, R., van Heult, P.T. and Xuan, J.A.  
TITLE Dna encoding a novel prost-ets polypeptide  
JOURNAL Patent: WO 0142472-A 10 14-JUN-2001;  
SCHERING AKTIENGESellschaft (DE)  
FEATURES  
source Location/Qualifiers  
1..23  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="antisense"

BASE COUNT 8 a 7 c 6 g 2 t  
ORIGIN

Query Match 1.2%; Score 23; DB 6; Length 23;  
Best Local Similarity 100.0%; Pred. No. 5.4e+06;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1734 tgactgacaagggcacagga 1756  
|||||  
DB 1 TGACTGCACAAAGGCCACAGGA 23

RESULT 6

AX175526  
LOCUS AX175526 23 bp mRNA linear PAT 03-JUL-2001  
DEFINITION Sequence 11 from Patent WO0142472.  
ACCESSION AX175526  
VERSION AX175526.1 GI:14598862  
KEYWORDS  
SOURCE  
ORGANISM synthetic construct.  
REFERENCE  
AUTHORS 1 (bases 1 to 23)  
Harkins, R., Lin, R.J., Luke, M., Monteciarlo, F., Parkes, D., Parry, G.,  
Stehlebrecher, R., van Heult, P.T. and Xuan, J.A.  
TITLE Dna encoding a novel pro-tecs polypeptide  
JOURNAL Patent: WO 0142472-A 11 14-JUN-2001;  
SCHERING AKTIEGESELSCHAFT (DE)  
FEATURES  
source 1..23  
Location/Qualifiers  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="antisense"  
BASE COUNT 6 a 8 c 5 g 4 t  
ORIGIN

Query Match 1.2%; Score 23; DB 6; Length 23;  
Best Local Similarity 100.0%; Pred. No. 5.4e+06;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1699 ccagcattccagagcagcct 1721  
|||||  
Db 1 CCAGCATTTCAGAGCAGAGCCT 23

RESULT 7  
AX199688  
LOCUS AX199688 50 bp DNA linear PAT 29-AUG-2001  
DEFINITION Sequence 618 from Patent WO0151670.  
ACCESSION AX199688  
VERSION AX199688.1 GI:15390129  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 50)  
Shimkets, R.A. and Leach, M.D.  
TITLE Nucleic acids containing single nucleotide polymorphisms and  
methods of use thereof  
JOURNAL Patent: WO 0151670-A 618 19-JUL-2001;  
Curagen Corporation (US)  
FEATURES  
source 1..50  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
misc\_feature 25..26  
/note="Nucleotide deleted between bases 25 and 26"  
Accession number cg44009326  
misc\_feature 26  
/note="2 of 2 allelic variants (617 is other entry)"  
BASE COUNT 4 a 12 c 25 g 9 t  
ORIGIN

Query Match 1.2%; Score 22.8; DB 6; Length 50;  
Best Local Similarity 66.0%; Pred. No. 4.8e+06;  
Matches 33; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 485 tcggagcagcgttgagcgagcgagcgagcgttggtcgcagag-534  
|||||  
Db 1 TCggGcCAGGCTGGCGCGAGGCGCGGCTGTGTTGTTACGTCAg 50

RESULT 8

AX221945  
LOCUS AX221945 48 bp mRNA linear PAT 07-SEP-2001  
DEFINITION Sequence 7387 from Patent WO0159103.  
ACCESSION AX221945  
VERSION AX221945.1 GI:15549669  
KEYWORDS  
SOURCE  
ORGANISM synthetic construct.  
REFERENCE  
AUTHORS 1 (bases 1 to 48)  
Blatt, L., McSwigen, J. and Chowrira, B.M.  
TITLE Method and reagent for the modulation and diagnosis of cd20 and  
nogo gene expression  
JOURNAL Patent: WO 0159103-A 7387 16-AUG-2001;  
RIBOZYME PHARMACEUTICALS, INC. (US); Blatt, Lawrence (US);  
McSwigen, James (US); Chowrira, Bharat M. (US)  
FEATURES  
source 1..48  
Location/Qualifiers  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Nucleic Acid"  
BASE COUNT 10 a 13 c 18 g 7 t  
ORIGIN

Query Match 1.2%; Score 22.6; DB 6; Length 48;  
Best Local Similarity 68.9%; Pred. No. 5.3e+06;  
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 137 ggcgtgggttgaggagctccctacagcagcagccctgagaccg 181  
|||||  
Db 1 GGCTGGCGGAGCAACCTCTCAAGACATCTCGGGAGCTG 45

RESULT 9  
AX249190  
LOCUS AX249190 31 bp DNA linear PAT 28-SEP-2001  
DEFINITION Sequence 1269 from Patent WO0166800.  
ACCESSION AX249190  
VERSION AX249190.1 GI:15863813  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 31)  
Cargill, M., Ireland, J.S. and Lander, E.S.  
TITLE Human single nucleotide polymorphisms  
JOURNAL Patent: WO 0166800-A 1269 13-SEP-2001;  
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)  
FEATURES  
source 1..31  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 5 a 11 c 6 g 8 t 1 others  
ORIGIN

Query Match 1.1%; Score 21.6; DB 6; Length 31;  
Best Local Similarity 80.0%; Pred. No. 8.6e+06;  
Matches 24; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 59 ccagctcccaagcctgtccagctccct 88  
|||||  
Db 1 CCAGCTCCCAAGCCYGGGTGATCTTCT 30

RESULT 10  
MUSJHCD6/c  
LOCUS MUSJHCD6 49 bp DNA linear ROD 27-APR-1993  
DEFINITION Mouse JH-cdelta locus sequence.  
ACCESSION M64562  
VERSION M64562.1 GI:198478

KEYWORDS C-region; Immunoglobulin; Immunoglobulin heavy chain; switch recombination junction.  
SOURCE Mus musculus (strain BALB/c, sub-species domesticus) DNA.  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 49)  
AUTHORS Owens,J.D., Jr., Finkelstein,F.D., Mountz,J.D. and Mushinski,J.F.  
TITLE Nonhomologous recombination at sites within the mouse JH-Cdelta locus accompanies Cmu deletion and switch to immunoglobulin D secretion  
JOURNAL Mol Cell. Biol. 11, 5660-5670 (1991)  
MEDLINE 92017847  
COMMENT This reference reports bases 1 to 25.  
FEATURES  
source Location/Qualifiers  
1. 49  
/organism="Mus musculus"  
/strain="BALB/C"  
/sub\_species="domesticus"  
/db\_xref="taxon:10090"  
/cell\_line="KWD8"  
/cell\_type="Hybridoma"  
BASE COUNT 7 a 17 c 14 g 11 t  
ORIGIN

Query Match 1.1%; Score 21.4; DB 10; Length 49;  
Best Local Similarity 66.0%; Pred. No. 8.3e+06;  
Matches 31; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 277 cccagatgctgctgagacagccagtcgctcagctccacacc 323  
Db 49 CCAGATGCTGGCTGGACAGACAGCTGCTGCTGCCACCC 3

RESULT 11  
A05116 A05116 45 bp DNA linear PAT 04-MAY-1993  
LOCUS Oligonucleotide U7 for bovine growth hormone.  
DEFINITION A05116  
ACCESSION A05116  
VERSION A05116.1 GI:345003  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 45)  
TITLE METHOD FOR THE PRODUCTION OF BOVINE GROWTH HORMONE USING A SYNTHETIC GENE  
JOURNAL Patent: WO 8805078-A 13 14-JUL-1988;  
FEATURES  
source Location/Qualifiers  
1. 45  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
BASE COUNT 8 a 17 c 10 g 10 t  
ORIGIN

Query Match 1.1%; Score 21.2; DB 6; Length 45;  
Best Local Similarity 76.5%; Pred. No. 9.1e+06;  
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1691 ctctgaccagcatctccagagacagagcctaca 1724  
Db 1 CTCGGGCCCTGCAGTTCCTCAGCAGAGTCTTCA 34

RESULT 12  
HSTCK6X45 48 bp mRNA linear PRI 05-FEB-1997  
LOCUS H.sapiens mRNA for T cell receptor beta chain junctional region  
DEFINITION (clone K6-45).  
ACCESSION 269575

VERSION 269575.1 GI:1770552  
KEYWORDS diversity region; joining region; T cell receptor; variable region.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 48)  
AUTHORS Naserke,H.E., Durinovic-Bello,I., Seidel,D. and Ziegler,A.G.  
TITLE The T-cell receptor beta chain CDR3 region of BV8S1/BJ1S5 transcripts in type 1 diabetes  
JOURNAL Immunogenetics 45 (2), 87-96 (1996)  
MEDLINE 97113296  
REFERENCE 2 (bases 1 to 48)  
AUTHORS Naserke,H.E.  
TITLE Direct Submission  
JOURNAL Submitted (02-FEB-1996) Naserke H.E., Diabetes Research Institute, Koelner Platz 1, D-80804 Muenchen, FRG  
FEATURES  
source Location/Qualifiers  
1. 48  
/organism="Homo sapiens"  
/isolate="healthy control WS"  
/db\_xref="taxon:9606"  
/tissue\_type="peripheral blood"  
/cell\_type="lymphocytes"  
/clone="K6-45"  
/note="HLA type: HLA-A31, 1; Cw3, 3; B62, 55; DRB1\*1301, 0401; DRB3\*0202, -; DQA1\*0103, 0301; DQB1\*0603, 0302; DPA1\*-; 01; DPB1\*0301, 0201"  
BASE COUNT 8 a 13 c 16 g 11 t  
ORIGIN

Query Match 1.1%; Score 21.2; DB 9; Length 48;  
Best Local Similarity 69.0%; Pred. No. 9e+06;  
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 651 gggccagcagtcggagagaccacctgaccagcctgagcagt 692  
Db 2 GTGCCAGCAGCGCGGGCGGTAAATGATCAACGCCACACATT 43

RESULT 13  
HSTRK6X45 48 bp mRNA linear PRI 05-FEB-1997  
LOCUS H.sapiens mRNA for T cell receptor beta chain junctional region  
DEFINITION (clone K6-45).  
ACCESSION 269547  
VERSION 269547.1 GI:1770714  
KEYWORDS diversity region; joining region; T cell receptor; variable region.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 48)  
AUTHORS Naserke,H.E., Durinovic-Bello,I., Seidel,D. and Ziegler,A.G.  
TITLE The T-cell receptor beta chain CDR3 region of BV8S1/BJ1S5 transcripts in type 1 diabetes  
JOURNAL Immunogenetics 45 (2), 87-96 (1996)  
MEDLINE 97113296  
REFERENCE 2 (bases 1 to 48)  
AUTHORS Naserke,H.E.  
TITLE Direct Submission  
JOURNAL Submitted (02-FEB-1996) Naserke H.E., Diabetes Research Institute,

FEATURES	
source	Koelner Platz 1, D-80804 Muenchen, FRG
location/Qualifiers	
1. 48	
/organism="Homo sapiens"	
/isolate="healthy control WS"	
/db_xref="taxon:9606"	
/tissue_type="peripheral blood"	
/cell_type="lymphocytes"	
/clone="K6-45"	
/note="HLA type: HLA-A*31, 1; C*3, 3; B*2, 55; DRB1*1301, 0401; DRB3*0202, -; DQA1*0103, 0301; DQB1*0603, 0302; DPA1*- , 01; DPB1*0301, 0201"	
1. 11	
/gene="V beta 8.1"	
<1. 11	
/gene="V beta 8.1"	
12. 28	
/note="N-D-N region"	
29. .>48	
/gene="J beta 1.5"	
29. .48	
/gene="J beta 1.5"	
BASE COUNT	8 a 13 c 16 g 11 t
ORIGIN	

Query Match	1.1%;	Score 21.2;	DB 9;	Length 48;
Best Local Similarity	69.0%;	Pred. No. 9e+06;		
Matches	29;	Conservative	0;	Mismatches 13; Indels 0; Gaps 0;
Qy	651	ggggccagcacttggggagagccacccctgaccacgcgcagcact	692	
Db	2	gtgccacacacccggcgccgctttatgtgatgacaccccccacatt	43	

RESULT	14				
AX250311/c					
LOCUS	AX250311	50 bp	DNA	linear	PAT 05-OCT-2001
DEFINITION	Sequence 51 from Patent WO0168867.				
ACCESSION	AX250311				
VERSION	AX250311.1	GI:15984079			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	Leadlay, P.F., Staunton, J. and Oliyink, M.C.				
TITLE	Polyketides and their synthesis				
JOURNAL	Patent: WO 0168867-A 51 20-SEP-2001;				
	Biotica Technology Limited (GB)				
FEATURES					
	Location/Qualifiers				

Query Match	1.1%	Score 21.2;	DB 6;	length 50;
Best Local Similarity	76.5%	Pred. No. 8.9e+06;		
Matches	26;	Conservative	0;	Mismatches 8; Indels 0; Gaps 0;
Qy	27	gcaactctgcgcgtgtccacactgtccccaagacc	60	
Db	43	GGACTGCGCCGCTGTGCACACTGACCGACGGGCC	10	

RESULT_15			
AX175518			
LOCUS	AM175518	21 bp	DNA
DEFINITION	Sequence 3 from Patent WO0142472.		linear
ACCESSION	AX175518		PAT 03-JUL-2001

VERSION	AX175518.1	GI:14598857
KEYWORDS		
SOURCE	synthetic construct.	
ORGANISM	synthetic construct.	
REFERENCE	artificial sequence.	
AUTHORS	1 (bases 1 to 21)	
TITLE	Harkins,R., Lin,R.J., Luke,M., Montecclaro,F., Parkes,D., Parry,G.,	
JOURNAL	Steinbrecher,R., van Heuvel,P.T. and Xuan,D.A.	
FEATURES	Patent: WO 012472-A 3 14-JUN-2001;	
source	SCHERING AKTIENGESELLSCHAFT (DE)	
	Location/Qualifiers	
	1..21	
	/organism="synthetic construct"	
	/db_xref="taxon:37630"	
	/note="primer"	
BASE COUNT	8 a	6 c 5 g 2 t
ORIGIN		

```

Query Match      1.1%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 1732 agtgcactcgcaaaagccaca 1752
    |||||
Db   1 AGTGAAGTCGCAAAAGCCACA 21

```

Search completed: August 6, 2002, 22:33:25  
Job time: 9604 sec

Page 5





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 20:55:46 ; Search time 311.92 Seconds  
(without alignments)  
10425.224 Million cell updates/sec

Title: US-09-841-963a-1

Perfect score: 1894  
Sequence: 1 gctgactctccctccagcacc.....ataaagatracagagaactg 1894

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 1900032

Minimum DB seq length: 10  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	34	1.8	34	22 AAC83265	GSEF specific PCR
C 2	31	1.6	31	22 AAC83264	GSEF specific PCR
C 3	24.6	1.3	42	20 AAX5838	Antisense oligonuc
C 4	23.6	1.2	47	21 AA25828	Human map-related
C 5	23.6	1.2	50	22 AAL34612	Human SNP oligonuc
C 6	23	1.2	23	22 AAD07340	Human transcriptio
C 7	23	1.2	23	22 AAD07341	Antisense oligo #1
C 8	23	1.2	23	22 AAD07342	Antisense oligo #2
C 9	23	1.2	23	22 AAD07343	Antisense oligo #3

C 10	22.8	1.2	50	22 AAL31921	Human SNP oligonuc
C 11	22.8	1.2	50	22 AAH89837	Human coding sequ
C 12	22.6	1.2	48	23 ABK07387	Human NOGO Ambery
C 13	22.4	1.2	50	22 AAL28314	Human SNP oligonuc
C 14	22.4	1.2	50	22 AAL34308	Human SNP oligonuc
C 15	22	1.2	45	22 AAD19606	Mutagenic PCR prim
C 16	22	1.2	45	22 AAD19607	Mutagenic PCR prim
C 17	22	1.2	45	22 AAF74532	Human MKK3 site-di
C 18	22	1.2	45	22 AAF74533	Human MKK3 site-di
C 19	22	1.2	45	22 AAC64157	Human MKK mutageni
C 20	22	1.2	45	22 AAC64158	Hepatoma AS-30D Ty
C 21	21.8	1.2	49	18 AAT80517	Human SNP oligonuc
C 22	21.8	1.2	50	22 AAL32325	PDGF-β receptor bl
C 23	21.6	1.1	45	22 AAF60332	Probe #4 complemen
C 24	21.4	1.1	39	12 AA011438	Human Kv6.2 DNA fr
C 25	21.2	1.1	36	20 AAT23808	Human SNP oligonuc
C 26	21.2	1.1	50	22 AAL29118	Human transcriptio
C 27	21	1.1	21	22 AAD07338	Human SNP oligonuc
C 28	21	1.1	37	22 AAF56968	FIV gene cloning r
C 29	21	1.1	50	22 AAL34241	Human SNP oligonuc
C 30	20.8	1.1	29	21 AAA04697	Polymorphic fragme
C 31	20.8	1.1	48	23 ABR06858	Human NOGO Ambery
C 32	20.8	1.1	48	23 ABR06859	Human NOGO Ambery
C 33	20.8	1.1	49	23 ABA10736	Tail adaptor oligo
C 34	20.8	1.1	50	17 AAT40111	Human Mer receptor
C 35	20.8	1.1	50	19 AAV39429	Humanised anti-HM1
C 36	20.8	1.1	50	19 AAV39430	Humanised anti-HM1
C 37	20.8	1.1	50	20 AAX59472	Primer used in con
C 38	20.8	1.1	50	20 AAX59473	Human SNP involvin
C 39	20.8	1.1	50	23 ABL01057	Human NOGO Ambery
C 40	20.6	1.1	48	23 ABR06870	Human single nucle
C 41	20.4	1.1	31	22 AAL30781	Human/mouse histon
C 42	20.4	1.1	45	21 AAC81266	Human map-related
C 43	20.4	1.1	47	21 AAZ65894	Human NOGO Ambery
C 44	20.4	1.1	48	23 ABR07085	Human NOGO Ambery
C 45	20.4	1.1	48	23 ABR07086	Human NOGO Ambery

## ALIGNMENTS

RESULT 1	
ID AAC83265/C	standard; DNA; 34 BP.
XX AAC83265;	
XX	
XX	16-MAR-2001 (first entry)
XX	GSEF specific PCR primer SEQ ID 11.
DE	
XX	Transcription factor; gland-specific Ets transcription factor; GSEF;
KW	metastatic potential; cancer; tumour; metastasis; breast; prostate;
KW	leukaemia; lymphoma; sarcoma; melanoma; chromosome 6p21.1-6p21.3;
KW	PCR primer; ss.
XX	
XX	Homo sapiens.
OS	
XX	
PN	WO200070092-A1.
XX	
PD	23-NOV-2000.
XX	
XX	12-MAY-2000; 2000WO-US31173.
PF	
XX	14-MAY-1999; 99US-0134112.
PR	
XX	(CHIR) CHIRON CORP.
PA	
XX	Kaufmann J, Xin H, Harrowe G;
PI	
XX	WPI; 2001-041019/05.
XX	Detecting metastatic and potential metastatic cancerous cells, useful
PT	

```
PT for diagnosing, prognosing, grading and staging of cancers by detecting
PT gland-specific Ets transcription factor gene product in a biological
PT sample from a cell
XX
PS Example 7: Page 56; 95pp; English.
CC This invention relates to a method for the detection and determination of
CC the metastatic potential of a cell. The method comprises detecting a
CC gland-specific Ets transcription factor (GSEF) gene product in a test
CC sample. Detection of a GSEF gene product in the test sample in amount
CC lower than that in a normal cell, is indicative of a cell with high
CC metastatic potential. The method is useful for determining the metastatic
CC potential of a cell, for the diagnosis and prognosis of cancer as well as
CC grading and staging of cancers by detecting GSEF expression in a
CC biological test sample. The method may also be used to monitor patients
CC having a predisposition to develop a particular cancer. GSEF polypeptides
CC are useful for producing antibodies, in cancer diagnosis, prognosis,
CC grading, staging and management of breast and prostate tumours, and in
CC detecting polymorphisms in the sequence. GSEF genes and proteins are also
CC useful in gene therapy. GSEF gene product expression levels can be used
CC in conjunction with any tissue in which an alteration in GSEF gene
CC product expression levels is associated with development of a
CC cancer-associated phenotype. Cancers, which can be monitored include
CC cancers of the prostate, cervix, lung and colon, melanomas, colorectal
CC adenocarcinomas, Wilms' tumour, retinoblastoma, sarcomas, myosarcomas,
CC lung carcinomas, leukaemia, and lymphomas. The GSEF gene is located on
CC human chromosome 6, specifically at 6p21.1-6p21.3. The present sequence
CC represents a PCR primer specific for the human GSEF DNA sequence. The
CC primer is used for the chromosomal mapping of the GSEF gene.
XX
SQ Sequence 34 BP; 5 A; 10 C; 7 G; 12 T; 0 other;

Query Match          1.8%; Score 34; DB 22; Length 34;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1850 ccaggagtgcccccggaggaatgataaagata 1883
    |||||||
DB 34 CCAAGAGTGTCCCGCGGAGATGATATTAAGATA 1

RESULT 2
AAC83264
ID AAC83264 standard; DNA; 31 BP.
XX
AC AAC83264;
XX
DT 16-MAR-2001 (first entry)
XX
DE GSEF specific PCR primer SEQ ID 10.
XX
KW Transcription factor; gland-specific Ets transcription factor; GSEF;
KW metastatic potential; cancer; tumour; metastasis; breast; prostate;
KW leukemia; lymphoma; sarcoma; melanoma; chromosome 6p21.1-6p21.3;
XX
OS Homo sapiens.
XX
PN WO200070092-A1.
XX
PD 23-NOV-2000.
XX
PF 12-MAY-2000; 2000WO-US13173.
XX
PR 14-MAY-1999; 99US-0134112.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Kaufmann J, Xin H, Hartowe G;
XX
DR WPI; 2001-041019/05.
XX
```

```
PT Detecting metastatic and potential metastatic cancerous cells, useful
PT for diagnosing, prognosing, grading and staging of cancers by detecting
PT gland-specific Ets transcription factor gene product in a biological
PT sample from a cell
XX
PS Example 7: Page 56; 95pp; English.
CC This invention relates to a method for the detection and determination of
CC the metastatic potential of a cell. The method comprises detecting a
CC gland-specific Ets transcription factor (GSEF) gene product in a test
CC sample. Detection of a GSEF gene product in the test sample in amount
CC lower than that in a normal cell, is indicative of a cell with high
CC metastatic potential. The method is useful for determining the metastatic
CC potential of a cell, for the diagnosis and prognosis of cancer as well as
CC grading and staging of cancers by detecting GSEF expression in a
CC biological test sample. The method may also be used to monitor patients
CC having a predisposition to develop a particular cancer. GSEF polypeptides
CC are useful for producing antibodies, in cancer diagnosis, prognosis,
CC grading, staging and management of breast and prostate tumours, and in
CC detecting polymorphisms in the sequence. GSEF genes and proteins are also
CC useful in gene therapy. GSEF gene product expression levels can be used
CC in conjunction with any tissue in which an alteration in GSEF gene
CC product expression levels is associated with development of a
CC cancer-associated phenotype. Cancers, which can be monitored include
CC cancers of the prostate, cervix, lung and colon, melanomas, colorectal
CC adenocarcinomas, Wilms' tumour, retinoblastoma, sarcomas, myosarcomas,
CC lung carcinomas, leukaemia, and lymphomas. The GSEF gene is located on
CC human chromosome 6, specifically at 6p21.1-6p21.3. The present sequence
CC represents a PCR primer specific for the human GSEF DNA sequence. The
CC primer is used for the chromosomal mapping of the GSEF gene.
XX
SQ Sequence 31 BP; 8 A; 9 C; 13 G; 1 T; 0 other;

Query Match          1.6%; Score 31; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1552 caggagggggaacacactgcccaggggga 1582
    |||||||
DB 1 caggagggggaacacactgcccaggggga 31

RESULT 3
AAK35839/c
ID AAK35839 standard; DNA; 42 BP.
XX
AC AAK35839;
XX
DT 14-JUL-1999 (first entry)
XX
DE Antisense oligonucleotide for mammalian ACE mRNA.
XX
KW Antisense oligonucleotide; mammalian; angiotensin converting enzyme;
KW ACE; hypertension; ss.
XX
OS Synthetic.
XX
PN WO9915643-A2.
XX
PD 01-APR-1999.
XX
PF 25-SEP-1998; 98WO-US20121.
XX
PR 25-SEP-1997; 97US-0059661.
XX
PA (UYFL ) UNIV FLORIDA.
XX
PI Mohuczy D, Phillips MI;
XX
DR WPI; 1999-276982/23.
XX
PT New antisense oligonucleotide compositions
XX
```

XX Claim 4; Page 16; 145pp; English.  
PS  
CC AAX35833-49 represent antisense oligonucleotides for mammalian  
CC angiotensin converting enzyme (ACE) mRNA. The antisense  
CC oligonucleotides can be used for treating or preventing  
CC hypertension. The antisense oligonucleotides can also be  
CC labeled and used for visualizing ACE mRNA in a cell, and for  
CC producing transgenic animals.  
XX  
SQ Sequence 42 BP; 1 A; 17 C; 14 G; 10 T; 0 other;

Query Match 1.3%; Score 24.6; DB 20; Length 42;  
Best Local Similarity 76.9%; Pred. No. 2.9e+04;  
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 392 gccgcagcccaaacagcagcgatggcgacgcccagc 430  
DB 39 GCGCGCAGCAGCAGCAGCAGCGCGCAGCAGCAGCAGC 1

## RESULT 4

AAZ65828  
ID AAZ65828 standard; DNA; 47 BP.

AC AAZ65828;

DT 11-SEP-2001 (first entry)

DE Human map-related diallelic marker SEQ ID NO:175.

XX Human genome; diallelic marker; high density disequilibrium map;  
XX genomic map; haplotype; phenotype; polymorphic base; genotyping;  
XX haplotyping; hybridisation; identification; characterisation;  
XX diagnosis; single nucleotide polymorphism; SNP; ds.  
OS Homo sapiens.

XX

FT Key Location/Qualifiers  
FT Variation replace(24,G)  
FT /tag= a  
FT /standard\_name= "single nucleotide polymorphism"

PN WO954500-A2.

PD 28-OCT-1999.

XX 21-APR-1999; 99WO-IB00822.

XX 21-APR-1998; 98US-0082614.

PR 23-NOV-1998; 98US-0109732.

XX (GEST) GENSET.

PI Cohen D, Blumenfeld M, Chumakov I;

XX WPI; 2000-013267/01.

PT Novel diallelic markers used to construct a high density disequilibrium  
PT map of the human genome -

XX Claim 1; Page 267; 2745pp; English.

XX AAZ65654 to AAZ69578 represent human diallelic markers from the present  
CC invention, which contain a polymorphic base at position 24 of their  
CC nucleotide sequences. AAZ65654 to AAZ77440 represent amplification  
CC primers for the diallelic markers. The diallelic markers of the  
CC invention have a variety of uses: they can be used for high density  
CC mapping of the human genome, and in complex association studies and  
CC haplotyping studies which are useful in determining the genetic basis  
CC for disease states. Compositions and methods of the invention can also  
CC be useful for the identification of the targets for the development of

CC pharmaceutical agents and diagnostic methods, as well as the  
CC characterisation of the differential efficacious responses to and side  
CC effects from pharmaceutical agents acting on a disease as well as other  
CC treatment.  
CC N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297  
CC and 3367, are not actually given a sequence in the sequence listing  
CC from the present invention.  
XX

SQ Sequence 47 BP; 6 A; 14 C; 16 G; 11 T; 0 other;

Query Match 1.2%; Score 23.6; DB 21; Length 47;  
Best Local Similarity 69.6%; Pred. No. 4.8e+04;  
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 1484 agccagccctcgatggtgggaaacggcgatgtgtctgtctct 1529  
DB 1 aggcagccctgtaccctcggaagcggctgtgtgtctgtctgt 46

## RESULT 5

AAI34612/C  
ID AAI34612 standard; DNA; 50 BP.

AC AAI34612;

DT 24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #7820.

XX Immunosuppressive; immunostimulatory; antineoplastic; cytostatic;  
XX immunoprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;  
XX amyloid protein; angiotensin; apoptosis related protein; cadherin;  
XX cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
XX complement related protein; cytochrome; kinase; cytokine; interferon;  
XX interleukin; G-protein coupled receptor; thioesterase; inflammation;  
XX multifactorial disease; autoimmune disease; infection;  
XX nervous system disease; ss.

XX Homo sapiens.

XX WO200147944-A2.

PN 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.

XX 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

XX (CUBA-) CUBAGEN CORP.

PI Shinkets RA, Leach M;

XX WPI; 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
XX oncogenes and histones, useful for diagnosing and treating, e.g.  
XX cancer, autoimmune diseases and infections -

PT Claim 1; Page 3652; 4143pp; English.

XX The present invention relates to oligonucleotides encoding polymorphic  
CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
CC histones, kinases, colony stimulating factors, complement related  
CC proteins, cytochromes, kinases, cytokines, interferons, interleukins,  
CC G-protein coupled receptors and thioesterases. The present sequence is  
CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
CC by them may be used in the prevention, diagnosis and treatment of  
CC diseases associated with inappropriate expression of the proteins listed  
CC above. Disorders that may be prevented, diagnosed and/or treated include  
CC multifactorial diseases with a genetic component, such as autoimmune

CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
 CC leukemia), diseases of the nervous system and an infection of pathogenic  
 CC organisms.

XX  
 SO Sequence 50 BP; 8 A; 17 C; 18 G; 7 T; 0 other;

#### Query Match

Best Local Similarity 1.2%; Score 23.6; DB 22; Length 50;  
 Best Local Similarity 69.6%; Pred. No. 4.8e+04;  
 Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 273 gcccccagatcgtgctgacagacacgcagctgacgtgccc 318  
 |||||  
 Db 47 cccccccagcgcgtgcacgtgccacggaagtcctcacagctcc 2

#### RESULT 6

AAD07340  
 ID AAD07340 standard; DNA; 23 BP.

XX  
 AC AAD07340;

DT 04-AUG-2001 (first entry)

XX  
 DE Human transcription factor, prost-ets mRNA expressing Tagman probe.

XX  
 KW Human: transcription factor protein; PROST-Ets; cytostatic; gene therapy;

KW vaccine; cancer; prostate; breast; ovary; benign prostatic hyperplasia;  
 KW antisense therapy; ribozyme therapy; therapeutic; cell proliferation; ss.

OS Homo sapiens.

XX  
 FH Key Location/Qualifiers

FT modified\_base 1

FT /tag- a

FT /mod\_base= OTHER

FT /note= "FAM-labeled Cytosine"

FT modified\_base 23

FT /tag- a

FT /mod\_base= OTHER

FT /note= "FAM-labeled Adenosine"

XX  
 PN MO200142472-A2.

XX  
 PD 14-JUN-2001.

XX  
 PE 29-NOV-2000; 2000MO-USA2295.

XX  
 PR 30-NOV-1999; 99US-0168182.

XX  
 PR 21-NOV-2000; 2000US-0718159.

XX  
 PA (SCHD ) SCHERING AG.

XX  
 PI Harkins R, Lin RJ, Luke M, Montecarlo F, Parkes D, Parry G;

XX  
 PI Steinbrecher R, Van Heult PT, Xuan J;

XX  
 DR WPI; 2001-381692/40.

XX  
 PT New transcription factor polypeptides, designated PROST-Ets

XX  
 PT polypeptides, useful for diagnostic, research and therapeutic

XX  
 PS applications -

XX  
 PS Example 2; Page 48; 74pp; English.

XX  
 CC The invention relates to human transcription factor protein, designated

CC as PROST-Ets and its corresponding cDNA molecule. PROST-Ets proteins and  
 CC nucleic acids are useful for treating a disease-state in human patient,  
 CC such as prostate, breast and ovarian cancer which is associated with  
 CC inappropriate expression of PROST-Ets and where the patient is in need of  
 CC decreased levels or activity of PROST-Ets. PROST-Ets is useful for

CC diagnosing and treating diseases of cell proliferation, such as prostate  
 CC cancer and benign prostatic hyperplasia. Prost-ets nucleic acids are used  
 CC as DNA probes, as target for antisense and ribozyme therapy, as template  
 CC for producing antisense polynucleotides, and as diagnostic reagent for  
 CC research, biological, clinical and therapeutic purposes. Prost-ets DNA is  
 CC also used in gene therapy. PROST-Ets protein is useful for generating  
 CC antibodies, and for stimulating immune response in cells containing  
 CC PROST-Ets proteins. The present DNA sequence is Tagman probe which is  
 CC used for the expression of human transcription factor, prost-ets mRNA.

XX  
 SO Sequence 23 BP; 4 A; 12 C; 2 G; 5 T; 0 other;

#### Query Match

Best Local Similarity 1.2%; Score 23; DB 22; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 5.6e+04;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1783 cctgcctcccatctgcacacaca 1805  
 |||||  
 Db 1 cctgcctcccatctgcacacaca 23

#### RESULT 7

AAD07341  
 ID AAD07341 standard; RNA; 23 BP.

XX  
 AC AAD07341;

XX  
 DE 04-AUG-2001 (first entry)

XX  
 DE Antisense oligo #1 to evaluate human prost-ets in cell proliferation.

XX  
 KW Human: transcription factor protein; PROST-Ets; cytostatic; gene therapy;

KW vaccine; cancer; prostate; breast; ovary; benign prostatic hyperplasia;  
 KW antisense therapy; ribozyme therapy; therapeutic; cell proliferation; ss.

OS Homo sapiens.

XX  
 PN WO200142472-A2.

XX  
 PD 14-JUN-2001.

XX  
 PE 29-NOV-2000; 2000MO-USA2295.

XX  
 PR 30-NOV-1999; 99US-0168182.

XX  
 PR 21-NOV-2000; 2000US-0718159.

XX  
 PA (SCHD ) SCHERING AG.

XX  
 PI Harkins R, Lin RJ, Luke M, Montecarlo F, Parkes D, Parry G;

XX  
 PI Steinbrecher R, Van Heult PT, Xuan J;

XX  
 DR WPI; 2001-381692/40.

XX  
 PT New transcription factor polypeptides, designated PROST-Ets

XX  
 PT polypeptides, useful for diagnostic, research and therapeutic

XX  
 PS applications -

XX  
 PS Claim 36; Page 56; 74pp; English.

XX  
 CC The invention relates to human transcription factor protein, designated

CC as PROST-Ets and its corresponding cDNA molecule. PROST-Ets proteins and  
 CC nucleic acids are useful for treating a disease-state in human patient,  
 CC such as prostate, breast and ovarian cancer which is associated with  
 CC inappropriate expression of PROST-Ets and where the patient is in need of  
 CC decreased levels or activity of PROST-Ets. PROST-Ets is useful for

CC diagnosing and treating diseases of cell proliferation, such as prostate  
 CC cancer and benign prostatic hyperplasia. Prost-ets nucleic acids are used  
 CC as DNA probes, as target for antisense and ribozyme therapy, as template  
 CC for producing antisense polynucleotides, and as diagnostic reagent for  
 CC research, biological, clinical and therapeutic purposes. Prost-ets DNA is  
 CC also used in gene therapy. PROST-Ets protein is useful for generating  
 CC antibodies, and for stimulating immune response in cells containing

CC PROST-Ets proteins. The present DNA sequence is antisense oligonucleotide  
 CC which is used for evaluating the role of human transcription factor,  
 CC PROST-Ets in cell proliferation to reduce Prost-Ets mRNA expression in  
 CC the selected prostate tumour cell lines.  
 XX  
 XX  
 SQ Sequence 23 BP; 4 A; 8 C; 5 G; 6 U; 0 other;

Query Match 1.2%; Score 23; DB 22; Length 23;  
 Best Local Similarity 73.9%; Pred. No. 5.6e+04;  
 Matches 17; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1161 tccacctgtgcagctctcctcaag 1183  
 :|||||:|||||:|||||:|||||:  
 Db 1 uccaccugugcagcaguccuccaag 23

## RESULT 8

AAD07342  
 ID AAD07342 standard; RNA; 23 BP.

AC AAD07342;

DT 04-AUG-2001 (first entry)

DE Antisense oligo #2 to evaluate human prost-ets in cell proliferation.

OS Human; transcription factor protein; PROST-Ets; cytostatic; gene therapy;  
 KM vaccine; cancer; prostate; breast; ovarian; benign prostatic hyperplasia;  
 KW antisense therapy; ribozyme therapy; therapeutic; cell proliferation; ss.

OS Homo sapiens.

PN WO200142472-A2.

PD 14-JUN-2001.

PF 29-NOV-2000; 2000WO-US42295.

PR 30-NOV-1999; 99US-0168182.

PR 21-NOV-2000; 2000US-0718159.

PA (SCHD ) SCHERING AG.

PI Hartins R, Lin RJ, Luke M, Montecarlo F, Parkes D, Parry G;

PI Steinbrecher R, Van Heult PT, Xuan J;

DR WPI; 2001-381692/40.

PS New transcription factor polypeptides, designated PROST-Ets  
 PT polypeptides, useful for diagnostic, research and therapeutic  
 PT applications

PS Claim 37; Page 56; 74pp; English.

CC The invention relates to human transcription factor protein, designated  
 CC as PROST-Ets and its corresponding cDNA molecule. PROST-Ets proteins and  
 CC nucleic acids are useful for treating a disease-state in human patient,  
 CC such as prostate, breast and ovarian cancer which is associated with  
 CC inappropriate expression of PROST-Ets and where the patient is in need of  
 CC decreased levels or activity of PROST-Ets. PROST-Ets is useful for  
 CC diagnosing and treating diseases of cell proliferation, such as prostate  
 CC cancer and benign prostatic hyperplasia. PROST-Ets nucleic acids are used  
 CC as DNA probes, as target for antisense and ribozyme therapy, as template  
 CC for producing antisense polynucleotides, and as diagnostic reagent for  
 CC research, biological, clinical and therapeutic purposes. Prost-ets DNA is  
 CC also used in gene therapy. PROST-Ets protein is useful for generating  
 CC antibodies, and for stimulating immune response in cells containing  
 CC PROST-Ets proteins. The present DNA sequence is antisense oligonucleotide  
 CC which is used for evaluating the role of human transcription factor,  
 CC PROST-Ets in cell proliferation to reduce Prost-ets mRNA expression in  
 CC the selected prostate tumour cell lines.

SQ Sequence 23 BP; 8 A; 7 C; 6 G; 2 U; 0 other;

Query Match 1.2%; Score 23; DB 22; Length 23;  
 Best Local Similarity 91.3%; Pred. No. 5.6e+04;  
 Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1734 tgactgcagcaagccacagca 1756  
 :|||||:|||||:|||||:|||||:  
 Db 1 ugacucgacaaagccacagca 23

## RESULT 9

AAD07343  
 ID AAD07343 standard; RNA; 23 BP.

AC AAD07343;

DT 04-AUG-2001 (first entry)

DE Antisense oligo #3 to evaluate human prost-ets in cell proliferation.

OS Human; transcription factor protein; PROST-Ets; cytostatic; gene therapy;  
 KM vaccine; cancer; prostate; breast; ovarian; benign prostatic hyperplasia;  
 KW antisense therapy; ribozyme therapy; therapeutic; cell proliferation; ss.

OS Homo sapiens.

PN WO200142472-A2.

PD 14-JUN-2001.

PF 29-NOV-2000; 2000WO-US42295.

PR 30-NOV-1999; 99US-0168182.

PR 21-NOV-2000; 2000US-0718159.

PA (SCHD ) SCHERING AG.

PI Hartins R, Lin RJ, Luke M, Montecarlo F, Parkes D, Parry G;

PI Steinbrecher R, Van Heult PT, Xuan J;

DR WPI; 2001-381692/40.

PS New transcription factor polypeptides, designated PROST-Ets  
 PT polypeptides, useful for diagnostic, research and therapeutic  
 PT applications

PS Claim 38; Page 56; 74pp; English.

CC The invention relates to human transcription factor protein, designated  
 CC as PROST-Ets and its corresponding cDNA molecule. PROST-Ets proteins and  
 CC nucleic acids are useful for treating a disease-state in human patient,  
 CC such as prostate, breast and ovarian cancer which is associated with  
 CC inappropriate expression of PROST-Ets and where the patient is in need of  
 CC decreased levels or activity of PROST-Ets. PROST-Ets is useful for  
 CC diagnosing and treating diseases of cell proliferation, such as prostate  
 CC cancer and benign prostatic hyperplasia. PROST-Ets nucleic acids are used  
 CC as DNA probes, as target for antisense and ribozyme therapy, as template  
 CC for producing antisense polynucleotides, and as diagnostic reagent for  
 CC research, biological, clinical and therapeutic purposes. Prost-ets DNA is  
 CC also used in gene therapy. PROST-Ets protein is useful for generating  
 CC antibodies, and for stimulating immune response in cells containing  
 CC PROST-Ets proteins. The present DNA sequence is antisense oligonucleotide  
 CC which is used for evaluating the role of human transcription factor,  
 CC PROST-Ets in cell proliferation to reduce Prost-ets mRNA expression in  
 CC the selected prostate tumour cell lines.

SQ Sequence 23 BP; 6 A; 8 C; 5 G; 4 U; 0 other;

Query Match 1.2%; Score 23; DB 22; Length 23;  
 Best Local Similarity 82.6%; Pred. No. 5.6e+04;



DE Human NOGO Amberzyme substrate sequence #547.  
 XX Human; ss; antisense therapy; cytostatic; antiinflammatory; haemostatic;  
 KW cerebroprotective; neurotropic; neuroprotective; antiparkinsonian;  
 KW muscular; CD20; neurite growth inhibitor gene; NOGO; hammerhead ribozyme;  
 KW DNazyme; G-cleaver; amberzyme; zinzyme; lymphoma; leukaemia;  
 KW B-cell lymphoma; non-Hodgkin's lymphoma; NHL; lymphocytic leukaemia;  
 KW human immunodeficiency virus; HIV associated NHL; mantle-cell lymphoma;  
 KW MCL; immunocytoma; IMC; immune thrombocytopenia; stroke; dementia;  
 KW inflammatory arthropathy; central nervous system injury;  
 KW cerebrovascular accident; CVA; Alzheimer's disease; multiple sclerosis;  
 KW chemotherapy-induced neuropathy; amyotrophic lateral sclerosis; ALS;  
 KW Parkinson's disease; ataxia; Huntington's disease; substrate sequence;  
 KW Creutzfeldt-Jakob disease; muscular dystrophy; neurodegenerative disease.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200159103-A2.  
 PN 16-AUG-2001.  
 XX 09-FEB-2001; 2001MO-US04273.  
 PF 11-FEB-2000; 2000US-181797P.  
 XX 28-FEB-2000; 2000US-185516P.  
 PR 06-MAR-2000; 2000US-187128P.  
 XX (RIBO-) RIBOZYME PHARM INC.  
 PA (BLAT/) BLATT L.  
 PA (MCSM/) MCSMIRGEN J.  
 PA (CHOW/) CHOWRIGA B M.  
 XX Blatt L, MCSMigen J, Chowrira BM;  
 PI WPI; 2001-607195/69.  
 DR Nucleic acid molecules, e.g., enzymatic nucleic acids and antisense  
 XX constructs, which down regulate expression of a CD20 gene or neurite  
 PT growth inhibitor gene useful for treating, e.g., lymphoma, leukemia,  
 PT and central nervous system injury -  
 XX Claim 89; Page 130; 200pp; English.  
 XX The invention relates to a nucleic acid molecule which down regulates  
 CC expression of a CD20 gene and a nucleic acid molecule which down  
 CC regulates expression of a neurite growth inhibitor gene (NOGO).  
 CC The nucleic acids may be enzymatic nucleic acids (e.g. a ribozyme or a  
 CC DNazyme) an inozyme (an endolytic nucleic acid cleaving a an RNA molecule  
 CC possessing an NCH motif), a G-cleaver (cleaving RNA with a NYN  
 CC motif) or an amberzyme (cleaving RNA with an NGN triplet), a zinzyme  
 CC (cleaving RNA with a YGY motif). The CD20-targeting nucleic acid is used  
 CC to cleave RNA of CD20 in the presence of a divalent cation that is  
 CC preferably Mg<sup>2+</sup>. Furthermore, it may be contacted with a cell to reduce  
 CC CD20 activity of the cell and treat a patient having a condition  
 CC associated with the level of CD20. The treatment may further comprise the  
 CC use of one or more therapies. In particular, the CD20 targeting  
 CC nucleic acid may be used to treat lymphoma, leukaemia, B-cell  
 CC lymphoma, low-grade or follicular non-Hodgkin's lymphoma (NHL), bulky  
 CC low-grade or follicular NHL, lymphocytic leukaemia, HIV (human  
 CC immunodeficiency virus) associated NHL, mantle-cell lymphoma (MCL),  
 CC immunocytoma (IMC), small B-cell lymphocytic lymphoma, immune  
 CC thrombocytopenia, and inflammatory arthropathy. The NOGO-targeting  
 CC nucleic acid is used to cleave RNA of the NOGO gene in the presence of a  
 CC divalent cation that is preferably Mg<sup>2+</sup>. Furthermore, the nucleic acid  
 CC may be contacted with a cell to reduce NOGO activity of the cell and  
 CC treat a patient having a condition associated with the level of NOGO. The  
 CC treatment may further comprise the use of one or more therapies.  
 CC In particular, the NOGO-targeting nucleic acid may be used to treat  
 CC central nervous system (CNS) injury and cerebrovascular accident (CVA,  
 CC stroke), Alzheimer's disease, dementia, multiple sclerosis (MS),  
 CC chemotherapy-induced neuropathy, amyotrophic lateral sclerosis (ALS),  
 CC Parkinson's disease, ataxia, Huntington's disease, Creutzfeldt-Jakob

CC disease, muscular dystrophy, and/or other neurodegenerative disease  
 CC states which respond to the modulation of NOGO expression. The  
 CC present sequence is a substrate sequence for a nucleic acid of the  
 CC invention based on the human NOGO sequence.  
 XX Sequence 48 BP; 10 A; 13 C; 18 G; 7 U; 0 other;  
 SO  
 Query Match 1.2%; Score 22.6; DB 23; Length 48;  
 Best Local Similarity 62.2%; Pred. No. 7.8e+04;  
 Matches 28; Conservative 3; Mismatches 14; Indels 0; Gaps 0;  
 OY 137 ggcgggggtagggaactccctacagcagcagccctgagaccg 181  
 |||:||||| |||:||||| |||:||||| |||:|||||  
 Db 1 ggcugggcgaggagaacuccucaaagcaugcuggggacug 45  
 RESULT 13  
 ID AAL28314  
 XX AAL28314 standard; DNA; 50 BP.  
 XX  
 AC AAL28314;  
 XX  
 DT 24-JAN-2002 (first entry)  
 DE  
 XX Human SNP oligonucleotide #1522.  
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;  
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
 KW multifactorial disease; autoimmune disease; infection;  
 KW nervous system disease; ss.  
 XX Homo sapiens.  
 OS  
 XX WO200147944-A2.  
 PN 05-JUL-2001.  
 XX 28-DEC-2000; 2000MO-US35498.  
 PF 28-DEC-1999; 99US-0173419.  
 XX 27-DEC-2000; 2000US-0173419.  
 XX (CURA-) CURAGEN CORP.  
 PA Shimkets RA, Leach M;  
 PI WPI; 2001-465210/50.  
 DR Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
 XX PT oncogenes and histones, useful for diagnosing and treating, e.g.  
 PT cancer, autoimmune diseases and infections -  
 XX Claim 1; Page 1817; 4143pp; English.  
 XX The present invention relates to oligonucleotides encoding polymorphic  
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
 CC histones, kinases, colony stimulating factors, complement related  
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,  
 CC G-protein coupled receptors and thioesterases. The present sequence is  
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
 CC by them may be used in the prevention, diagnosis and treatment of  
 CC diseases associated with inappropriate expression of the proteins listed  
 CC above. Disorders that may be prevented, diagnosed and/or treated include  
 CC multifactorial diseases with a genetic component, such as autoimmune  
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,

CC leukaemia), diseases of the nervous system and an infection of pathogenic  
CC organisms.  
XX  
SQ Sequence 50 BP; 11 A; 16 C; 16 G; 7 T; 0 other;  
  
Query Match 1.2%; Score 22.4; DB 22; Length 50;  
Best Local Similarity 81.2%; Pred. No. 8.7e+04;  
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
OY 349 tgcagcacagctccctctgagccagcagc 380  
||||| | ||||| ||||| ||||| |  
Db 12 tgcagggaaagagctccctctgagcactgaccc 43  
  
RESULT 14  
AAL34308/C  
ID AAL34308 standard; DNA; 50 BP.  
XX  
AC AAL34308;  
XX  
DT 24-JAN-2002 (first entry)  
XX  
DE Human SNP oligonucleotide #7516.  
XX  
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;  
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200147944-A2.  
XX  
PD 05-JUL-2001.  
XX  
PF 28-DEC-2000; 2000WO-US35498.  
XX  
PR 28-DEC-1999; 99US-0173419.  
PR 27-DEC-2000; 2000US-0173419.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Shimkets RA, Leach M;  
XX  
DR WPI: 2001-465210/50.  
XX  
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
PT oncogenes and histones, useful for diagnosing and treating, e.g.  
PT cancer, autoimmune diseases and infections -  
XX  
PS Claim 1; Page 3554; 4143pp; English.  
XX  
CC The present invention relates to oligonucleotides encoding polymorphic  
CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
CC histones, kinases, colony stimulating factors, complement related  
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,  
CC G-protein coupled receptors and thioesterases. The present sequence is  
CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
CC by them may be used in the prevention, diagnosis and treatment of  
CC diseases associated with inappropriate expression of the proteins listed  
CC above. Disorders that may be prevented, diagnosed and/or treated include  
CC multifactorial diseases with a genetic component, such as autoimmune  
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
CC (e.g. cancers of the bladder, brain, breast, colon and kidney  
CC leukaemia), diseases of the nervous system and an infection of pathogenic  
CC organisms.

XX  
SQ Sequence 50 BP; 3 A; 14 C; 19 G; 14 T; 0 other;  
  
Query Match 1.2%; Score 22.4; DB 22; Length 50;  
Best Local Similarity 66.7%; Pred. No. 8.7e+04;  
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
  
OY 1609 catggggcaggggtgctctcctcagccagctgctccctggagg 1656  
|| ||||| | ||||| | ||||| ||||| ||||| |||||  
Db 49 CAAGGGCCAGACTACTCCGACAGCCAGCCAGCCCAACCCCAAGG 2  
  
RESULT 15  
AAD19606  
ID AAD19606 standard; DNA; 45 BP.  
XX  
AC AAD19606;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Mutagenic PCR primer, SER-U to introduce mutation in human MKK3 gene.  
XX  
DE 1,3-thiazole compound; antiinflammatory; immunosuppressive; nocotropic;  
KW neuroprotective; immunomodulator; antiarteriosclerotic; dermatological;  
KW antibacterial; hypertensive; p38 MAP kinase; phosphodiesterase IV;  
KW PDE IV; virucide; tumour necrosis factor alpha; TNF; Addison's disease;  
KW cytokine-mediated disease; adenosine receptor-mediated disease; diabetes;  
KW autoimmune hemolytic anemia; Crohn's disease; psoriasis; rheumatism;  
KW spinal cord injury; multiple sclerosis; Alzheimer's disease; arthritis;  
KW Parkinson's disease; amyotrophic lateral sclerosis; ulcerative colitis;  
KW toxemia; chronic pneumonia; pulmonary silicosis; pulmonary sarcoidosis;  
KW lung tuberculosis; Creutzfeldt-Jakob disease; arterial sclerosis; AIDS;  
KW acquired immune deficiency syndrome; systemic lupus erythematosus;  
KW cachexia; atopic dermatitis; meningitis; encephalopathy; angina pectoris;  
KW myocardial infarction; congestive heart failure; hepatitis; transplant;  
KW dialysis hypotension; diffuse intravascular coagulation syndrome;  
KW human; MKK3 gene; PCR primer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200174811-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 29-MAR-2001; 2001WO-JP02629.  
XX  
PR 30-MAR-2000; 2000JP-0097876.  
PR 02-FEB-2001; 2001JP-0027571.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Ohkawa S, Naruo K, Miwatashi S, Kimura H;  
XX  
DR WPI: 2001-616682/71.  
XX  
PT New 1,3-thiazole compounds (I) or their salts (with position 5  
PT substituted by 4-pyridyl) for prevention and treatment of diseases  
PT including cytokine-mediated disease, adenosine receptor-mediated  
PT diseases or inflammation -  
XX  
PS Example 4; Page 263; 288pp; English.  
XX  
CC The present invention relates to 1,3-thiazole compounds or their  
CC salts (with position 5 substituted by 4-pyridyl group having a  
CC substituent including no aromatic group, at a position adjacent  
CC to the N of pyridyl group). The compounds of the invention are  
CC excellent inhibitors of p38 MAP kinase, phosphodiesterase IV (PDE  
CC IV) and tumour necrosis factor (TNF) alpha production. They are  
CC antagonists of adenosine receptor. They are used for the prevention  
CC and treatment of cytokine-mediated diseases, adenosine receptor-  
CC mediated diseases, inflammation, Addison's disease, autoimmune  
CC haemolytic anaemia, Crohn's disease, psoriasis, rheumatism, spinal



CC cord injury, multiple sclerosis Alzheimer's disease, Parkinson's  
CC disease, amorphotropic lateral sclerosis, diabetes, arthritis, toxemia  
CC ulcerative colitis, chronic pneumonia, pulmonary silicosis, pulmonary  
CC sarcoidosis, lung tuberculosis, cachexia, Creutzfeldt-Jakob disease,  
CC arterial sclerosis, virus infection, atopic dermatitis, systemic lupus  
CC erythematosus, meningitis, AIDS (acquired immune deficiency syndrome),  
CC encephalopathy, angina pectoris, myocardial infarction, congestive  
CC heart failure, hepatitis, transplant, dialysis hypotension or diffuse  
CC intravascular coagulation syndrome. The present DNA sequence is PCR  
CC primer, SER-v which is used for generating an active type of human  
CC MKK3 gene by introducing a mutation (S189E and T193E). MKK3 gene is  
CC inhibited by the compounds of the present invention.

Query Match	1.2%;	Score 22;	DB 22;	Length 45;
Best Local Similarity	73.7%;	Pred. No. 1e+05;		
Matches	28;	Conservative	0;	Mismatches 10;
			Indels	0;
			Gaps	0;

QY 792 t g t g t g g c g a a g t g t c t c a a g g a c a t c g a g a c g g c t g c 823  
||||| ||| ||| ||||| ||| | | |||||  
Db 8 t g t g t g a c g a g t g t g c c a a g a g a g a t g t g a t c g c g c t g c 45

```
Search completed: August 6, 2002, 22:40:31
Job time: 6285 sec
```

Job time: 6285 sec





```

? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-241-1300
? TELEFAX: 703-241-2848
? INFORMATION FOR SEQ ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 45 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA
? HYPOTHETICAL: NO
? FEATURE:
? NAME/KEY: -
? LOCATION: 1..45
? OTHER INFORMATION: /label=oligonucleotide
? OTHER INFORMATION: /note="UV oligonucleotide portion of synthetic
? OTHER INFORMATION: BGH gene, Figure 1."
? JS-07-863-689A-7

```

```
Qy      1691 cttgacccaccagcatlccagagcagaagcctaca 1724
          ||| ||||| ||| ||||| ||||| ||||| |||||
Db       1 CTCGGGCCCCCTGCAGTTCCTCAGACAGAGTCTTTCA 34
```

RESULT 4  
US-08-628-747-23/c  
; Sequence 23, Application US/08628747  
; Patent No. 6169070

```
;; GENERAL INFORMATION:
;;
;; APPLICANT:  Chen, Jian
;;
;; APPLICANT:  Godowski, Paul J.
```

APPLICANT: Mather, Jennie P.  
APPLICANT: Li, Ronghao  
TITLE OF INVENTION: RECEPTOR ACTIVATION BY GAS6  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: One DNA Way  
City: South San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/628,747  
FILING DATE: 17-APR-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/402,253  
FILING DATE: 10-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/438,861  
FILING DATE: 10-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: P929P2PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-225-1994  
TELEFAX: 650-952-9881  
TELEX: 910-371-7168  
INFORMATION FOR SEQ ID NO: 23:



;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: William M. Smith  
;; STREET: One Market Plaza, Stewart Tower, Suite 2000  
;; CITY: San Francisco  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94105  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/834,539A  
;; FILING DATE: 1992-02-05  
;;  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Smith, William M.  
;; REGISTRATION NUMBER: 30,223  
;; REFERENCE/DOCKET NUMBER: 14643-5  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-543-9600  
;; TELEFAX: 415-543-5043  
;; INFORMATION FOR SEQ ID NO: 50:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 42 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;;  
;; MOLECULE TYPE: DNA (genomic)  
;;  
US-07-834-539A-50

Query Match 1.1%; Score 20.2; DB 1; Length 42;  
Best Local Similarity 75.8%; Pred. No. 3.3e+04;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 625 cagcagctggcgagcccaagccctggggccag 657  
||||| || || |||| |||| || || ||||  
Db 1 CAGCAGCTGCACACCCCAATGCCATGAGCCCGAG 33

RESULT 8  
US-08-053-131-67  
; Sequence 67, Application US/08053131  
; Patent No. 5661016  
; GENERAL INFORMATION:  
; APPLICANT: Lonberg, Nils  
; TITLE OF INVENTION: Transgenic No. 5661016-Human Animals for  
; TITLE OF INVENTION: Producing Heterologous Antibodies  
; NUMBER OF SEQUENCES: 197  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourile and Crew  
; STREET: One Market Plaza, Stewart Tower, Suite 200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/053,131  
; FILING DATE: 26-APR-1993  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/990,860  
; FILING DATE: 16-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/810,279

;; FILING DATE: 17-DEC-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/853,408  
;; FILING DATE: 18-MAR-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Smith, William M.  
;; REGISTRATION NUMBER: 30,223  
;; REFERENCE/DOCKET NUMBER: 14643-9-3  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-326-2400  
;; TELEFAX: 415-326-2422  
;; INFORMATION FOR SEQ ID NO: 67:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 42 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;;  
;; MOLECULE TYPE: DNA (primer)  
;;  
US-08-053-131-67

Query Match 1.1%; Score 20.2; DB 1; Length 42;  
Best Local Similarity 75.8%; Pred. No. 3.3e+04;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 625 cagcagctggcgagcccaagccctggggccag 657  
||||| || || |||| |||| || || ||||  
Db 1 CAGCAGCTGCACACCCCAATGCCATGAGCCCGAG 33

RESULT 9  
US-08-645-641-67  
; Sequence 67, Application US/08645641  
; Patent No. 5719032  
; GENERAL INFORMATION:  
; APPLICANT: Lonberg, Nils  
; TITLE OF INVENTION: Transgenic No. 5719032-Human Animals for  
; TITLE OF INVENTION: Producing Heterologous Antibodies  
; NUMBER OF SEQUENCES: 150  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William M. Smith  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/645,641  
; FILING DATE: 20-MAY-1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/904,068  
; FILING DATE: 23-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 14643-000913  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (primer)



COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/800,353  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/834,539  
FILING DATE: 1992-02-05  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 14643-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-800-353-50

Query Match 1.1%; Score 20.2; DB 2; Length 42;  
Best Local Similarity 75.8%; Pred. No. 3.3e+04;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 625 cagcagctgggagcccaagccctggggccag 657  
||||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 1 CAGCAGGTGCACACCCATGCCCATGAGCCCG 33

RESULT 13  
US-08-308-865-67  
Sequence 67; Application US/08308865  
Patent No. 5877397  
GENERAL INFORMATION:  
APPLICANT: Lomborg, Nils  
APPLICANT: Kay, Robert M.  
TITLE OF INVENTION: Transgenic No. 5877397-Human Animals for  
TITLE OF INVENTION: Producing Heterologous Antibodies  
NUMBER OF SEQUENCES: 150  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: William M. Smith  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/308,865  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/145,707  
FILING DATE:  
APPLICATION NUMBER: US 07/904,068  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.

REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 14643-9-1-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (primer)  
US-08-308-865-67

Query Match 1.1%; Score 20.2; DB 2; Length 42;  
Best Local Similarity 75.8%; Pred. No. 3.3e+04;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 625 cagcagctgggagcccaagccctggggccag 657  
||||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 1 CAGCAGGTGCACACCCATGCCCATGAGCCCG 33

RESULT 14  
US-09-042-353-228  
Sequence 228; Application US/09042353  
Patent No. 6255458  
GENERAL INFORMATION:  
APPLICANT: Lomborg, Nils  
APPLICANT: Kay, Robert M.  
TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for  
TITLE OF INVENTION: Producing Heterologous Antibodies  
NUMBER OF SEQUENCES: 421  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/042,353  
FILING DATE: 13-MAR-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/810,279  
FILING DATE: 17-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/853,408  
FILING DATE: 18-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/904,068  
FILING DATE: 23-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/990,860  
FILING DATE: 16-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/053,131  
FILING DATE: 26-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/096,762  
FILING DATE: 22-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/155,301  
FILING DATE: 18-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/161,739







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 19:52:36 ; Search time 2035.51 Seconds  
(without alignments)  
12558.638 Million cell updates/sec

Title: US-09-841-963A-1

Sequence: 1 gctcgaacttcctccagcagac.....ataaagatactagagactg 1894

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 89406

Minimum DB seq length: 10  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estcl:\*  
10: gb\_estl2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_dln:\*  
16: em\_gss\_vrl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	1.6	47	10	R07222
2	26.2	1.4	40	9	AA789732
3	24.8	1.3	50	9	AU104281
4	24.8	1.3	50	9	AU104284
5	23.2	1.2	50	9	AU104283
6	22.6	1.2	49	9	AA108149
7	21.6	1.1	46	9	AI690532
8	21.4	1.1	50	9	AU104286
9	21.4	1.1	47	12	A2403867
10	21.4	1.1	50	9	AU104319
11	21	1.1	50	9	AU107649
12	21	1.1	50	9	AU107650
13	21	1.1	50	9	AU107651
14	20.8	1.1	50	9	AU104063
15	20.6	1.1	50	9	AU104502
16	20.4	1.1	48	10	BI462075
17	20.4	1.1	50	9	AU104329

C 18	20.4	1.1	50	9	AU105943	AU105943
C 19	20.2	1.1	50	9	AL586305	AL586305
C 20	20.2	1.1	50	9	AU102557	AU102557
C 21	20.2	1.1	50	9	AU104466	AU104466
C 22	20.2	1.1	50	9	AU104498	AU104498
C 23	20	1.1	44	12	A2783407	A2783407
C 24	19.8	1.0	46	9	AA632261	AA632261
C 25	19.8	1.0	49	9	AA986590	AA986590
C 26	19.8	1.0	50	9	AU104501	AU104501
C 27	19.8	1.0	50	9	AU104612	AU104612
C 28	19.6	1.0	36	12	A2807008	A2807008
C 29	19.6	1.0	43	10	BE373472	BE373472
C 30	19.6	1.0	46	10	BI488502	BI488502
C 31	19.6	1.0	49	9	A1355157	A1355157
C 32	19.6	1.0	49	10	BF571906	BF571906
C 33	19.4	1.0	43	9	A1570014	A1570014
C 34	19.4	1.0	47	9	AA290779	AA290779
C 35	19.4	1.0	48	12	BH634120	BH634120
C 36	19.4	1.0	50	9	AU102813	AU102813
C 37	19.4	1.0	50	9	AU104330	AU104330
C 38	19.4	1.0	50	9	AU105871	AU105871
C 39	19.4	1.0	50	10	BI256277	BI256277
C 40	19.2	1.0	42	10	BI823879	BI823879
C 41	19.2	1.0	49	9	AA677795	AA677795
C 42	19.2	1.0	49	9	AA781646	AA781646
C 43	19.2	1.0	50	9	AU104064	AU104064
C 44	19.2	1.0	50	9	AU104491	AU104491
C 45	19.2	1.0	50	9	AU104493	AU104493

#### ALIGNMENTS

RESULT 1  
LOCUS R07222  
DEFINITION yf1ab07.r1 Soares fetal liver spleen livers Homo sapiens cDNA clone IMAGE:126805 5' similar to gb:S45203 Flt-1 ONCOGENE (HUMAN);, mRNA sequence.

ACCESSION R07222  
VERSION R07222.1 GI:759145  
KEYWORDS EST.

SOURCE human.  
ORGANISM Homo sapiens  
Fukuroyuta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

#### AUTHORS

Unpublished (1995)  
The Washu-Merck EST Project

#### JOURNAL

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu

#### COMMENT

High quality sequence starts: 1 High quality sequence stops: 1  
Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Trace considered overall poor quality  
Insert length: 1185 Std Error: 0.00  
Seq primer: M13R1

High quality sequence stop: 1.  
Location/Qualifiers

#### FEATURES

source  
1..47  
/organism="Homo sapiens"  
/db\_xref="GDB:478966"  
/db\_xref="taxon:9606"  
/clone="IMAGE:126805"

```

/clone.lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACGCGAAGAAATTAATTAAGACCTTTTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

```

Query Match	1.68;	Score 30;	DB 10;	Length 47;
Best Local Similarity	78.3%;	Pred. No. 4.1e+04;		
Matches	36;	Conservative	0;	Mismatches 10;
			Indels	0;
			Gaps	0;

```

Oy      1314  tgaactgcgacaaagctgtagccgctccatccgcagctcttacagaagaa 1359
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1    TGAATTACGACAAGCTGAGCCGCTCGCTCCGATACTATTATGAGAA 46

```

LOCUS	DEFINITION	PROTEIN D-ETS-4
AA789732	40 bp mRNA	linear EST 06-FEB-1996
AA789732	v718602.r1 Barstead mouse irradiated colon MBLR81 Mus musculus cDNA clone IMAGE:1177251 5' similar to SW:ETS4_DROME P29775 DNA-BINDING	
		PROTEIN D-ETS-4 ; , mRNA sequence.

ACCESSION	AA789732	
VERSION	AA789732.1	GI:2849852
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (pages 1 to 40)

Matta, M., Hillier, L., Allen, M., Bowles, M., Diétrich, N., Dubouque, T.,

TITLE The WashU-RHMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Maira M/Mouse EST Project

Masnu-NHMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: [mousese@wustl.edu](mailto:mousese@wustl.edu)  
 This clone is available royalty-free through LNC; contact the  
 IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
 MGI:635099

FEATURES	source
Trace considered overall poor quality	
Possible reversed clone: similarity on wrong strand	
Seq primer: -28ml3 rev2 ET from AmerSham	
High quality sequence stop: 1.	
location/Qualifiers	
1..40	

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="1177251"
/clone_id="Barstead mouse irradiated colon MPLR37"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI. Tissue obtained
from 8 week old mouse. Colon was harvested 72 hours after

```

irradiation with 1400 Gys. 1st strand cDNA was primed with a Not I - oligo(dT) primer  
 5' TTTTCGGAATCTGAAAGTGGAGGCGGCCGCTTTTTTTTTTTTTTTTTTTT  
 T 3'; double-stranded cDNA was ligated to Eco RI  
 adaptors [AATTTCGAGCTCGG] digested with Not I and cloned  
 into the Not I and Eco RI sites of the modified pT773  
 vector. Library constructed by Bob Barsstead.

Query Match	1.48;	Score 26.2;	DB 9;	Length 40;
Best Local Similarity	79.58;	Pred. No. 2e+05;		
Matches 31; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;

```
Oy 1163 caactgtgycagttctctcaagaagttgtcctaactcaagccc 120
      || ||||| ||||| || || |||||
Db 2 CAACGTGTGGCAGTTCCTGAAAGAAGACTGGTGTCAAGCCC 40
```

RESULT	3						
AI104281/c							
LOCUS	AI104281	50 bp	MRNA	linear	EST 30-AUG-2001		
DEFINITION	AI104281	Sugano Homo sapiens	CDNA library	Homo sapiens	CDNA clone		
ACCESSION	HEP04739,	MRNA sequence.					
	AI104281						

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Eukaryota; Metazoa; Chordata; Carnalia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (pages 1 to 50)  
Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,I., Sese,J., Hatake,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.  
Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites  
EMBO Rep. 2 (3), 388-393 (2001)

MEDLINE  
COMMENT  
212/00/2  
Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp  
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, A. and Suganuma, S.  
Construction and characterization of a full length-enriched and  
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

```

FEATURES
  source
    1. location/Qualifiers
       1.50
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone="HEP04739"
       /clone_lib="Sugano Homo sapiens CDNA library"
BASE COUNT
  7 a 13 c 26 g 4 t
ORIGIN

```

Query Match	1.3%;	Score 24.8;	DB 9;	Length 50;
Best Local Similarity	80.6%;	Pred. No. 3.8e+05;		
Matches 29; Conservative	0;	Mismatches 7;	Indels 0;	Gaps 0.

**Oy**      457 ccacctcctgtctgcccccgaaacagtgctgcggagc 499  
          || ||||| |||| | ||||| ||||| ||  
**Db**      45 ccccttctctggcccgcgcgaaccgctgtgccgcgwc 10

RESULT	4
AU104284/c	
LOCUS	50 bp mRNA linear EST 30-AUG-2001
DEFINITION	Sugano Homo sapiens CDNA library Homo sapiens cDNA clone
HEB07220,	mRNA sequence.
ACCESSION	AU104284

VERSION	AU104284.1	GI:13553805
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 50)	
AUTHORS	Suzuki,Y., Tsira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Harae,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.	
TITLE	Diverse transcripional Initiation revealed by fine, large-scale mapping of mRNA start sites	
JOURNAL MEDLINE COMMENT	EMBO Rep. 2 (5), 388-393 (2001) Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: yusuzuki@ims.u-tokyo.ac.jp Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).	
FEATURES	source	
BASE COUNT	7 a 11 c 28 g 4 t	
ORIGIN	1..50 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="HEP07220" /clone_lib="Sugano Homo sapiens cDNA library"	
Query Match	1.3%; Score 24.8; DB 9;	
Best Local Similarity	80.6%; Pred. No. 3-Best;	
Matches	Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
OY	457 ccactctcgtcgccccccgacacgggtgcggagac 492	
DB	44 CCCCTCGGCGCCGCCGACCGCAGTCGTGGCGTC 9	
LOCUS	AU104283/C	
DEFINITION	AU104283 Sugano Homo sapiens cDNA library Homo sapiens CDNA clone HEF06519, mRNA sequence.	
ACCESSION	AU104283	
VERSION	AU104283.1	
KEYWORDS	GI:13553804 EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 50)	
AUTHORS	Suzuki,Y., Tsira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Harae,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.	
TITLE	Diverse transcripional Initiation revealed by fine, large-scale mapping of mRNA start sites	
JOURNAL MEDLINE COMMENT	EMBO Rep. 2 (5), 388-393 (2001) Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: yusuzuki@ims.u-tokyo.ac.jp Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).	
FEATURES	location/Qualifiers	
SOURCE	1..50 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="HEP07220" /clone_lib="Sugano Homo sapiens cDNA library"	

[illegible]



```

BASE COUNT
ORIGIN
      8 a
      13 c
      15 g
      11 t

/cloone.lib="Mouse 10kb plasmid UUCGCM library"
/sex="Male"
/lab.host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

Query Match	1.1%;	Score 21.4;	DB 16;	Length 47;
Best Local Similarity	71.8%;	Pred. No. 1.6e+02;		
Matches 28; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0;
Oy	1020	tgcagccacccctcgtgacatcttggaagtccagcgccctgta	1058	
Db	9	ttcatgcccacactggtgtctcttgagagggagacgactgtga	47	

RESULT	10
LOCUS	AI104319
DEFINITION	AI104319 50 bp mRNA linear EST 30-AUG-2003
ACCESSION	AI104319 Sgano Homo sapiens CDNA library Homo sapiens CDNA clone
VERSION	AI104319
KEYWORDS	HEP05158, mRNA sequence.
SOURCE	AI104319.1 GI:13553840
ORGANISM	EST.
	human.
	human.
	Homo sapiens
	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
1 (pages 1 to 50)	Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Gese,J., Hatae,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.	Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites	EMBO Rep. 2 (5), 388-393 (2001)	21270072	Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan Email: y Suzuki@iims.u-tokyo.ac.jp
	Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S.	Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).			Location/Qualifiers 1. .50

BASE COUNT	11 a	11 c	18 g	10 t
ORIGIN				

	Query Match	Similarity	Score	DB	Length
	Best Local	66.0%	Pred. No. 1,6e+06		
	Matches	31; Conservative	0; Mismatches	16; Indels	0; Gaps
Q7	778	ggttcagtcacatggtgtgtggcgaaagtcctcaagacacgtcaagcgg	824		
Db	4	gatggtgtccatctgtcgtccacgcgcgaattaataaagccattgcgccag	50		

RESULT	11
LOCUS	AU107649
DEFINITION	50 bp mRNA linear; ESN 30-AUG-2001
ACCESSION	AU107649
VERSION	AU107649.1
KEYWORDS	GI:13557170
	EST
	cdna library Homo sapiens
	cdna clone
	sequence.

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (Phases 1 to 50)
REFERENCE	Suzuki,T., Taira,H., Tsunoda,T., Miushima-Sugano,T., Sese,T., Hatae ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki ,Y., Nakamura,Y., Suyama,A. and Sugano,S. Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites EMBO Rep. 2 (5), 388-393 (2001)
JOURNAL	
COMMENT	Contact: Yutaka Suzuki 21270072

**FEATURES**  
location/qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HS102001"
/clone_1bp="Sugano Homo sapiens cDNA library"
BASE COUNT      17 a      11 c      15 g      7 t
ORIGIN

```

Query Match	1.18;	Score 21;	DB 9;	Length 50;
Best Local Similarity	82.88;	Pred. No. 1.9e+06;		
Matches 24; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;

```

Qy      862 ctgagagccccagcaatgtctcagaagtggc      890
          | | | | | | | | | | | | | | | |
Db      21  CCGAGCCCGAGCAGTGTGAAGAGAAGAGGC      49

```

RESULT	12				
AU107650					
LOCUS		50 bp	mRNA	linear	EST 30-AUG-2001
DEFINITION	AU107650	Sgungano Homo sapiens cDNA library	Homo sapiens cDNA clone		
ACCESSION	HS103379.	mRNA sequence.			
VERSION	AU107650				
KEYWORDS	AU107650.1	GI:13557171			
SOURCE	EST.				
	human.				

REFERENCE AUTHORS	TITLE
1 (bases 1 to 50)	
Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hatake,H., Ota,T., Isegai,T., Tanaka,T., Morishita,S., Okudo,R., Sakaki,Y., Nakamura,Y., Sugama,A. and Sugano,S.	Diverse transcriptional initiation revealed by fine, large-scale





Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: ysuzuki@ims.u-tokyo.ac.jp  
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano  
S. Construction and characterization of a full length-enriched and  
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

## FEATURES

## Source

1.50  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="ZRV6C934"  
/clone\_lib="Sugano Homo sapiens cDNA library"

## BASE COUNT

9 a 9 c 25 g 7 t

## ORIGIN

Query Match 1.1%; Score 20.6; DB 9; Length 50;  
Best Local Similarity 74.3%; Pred. No. 2.3e+06;  
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 501 agaaggcgagcgaggcgagtggtctcgagaga 535  
||| ||||| ||||| || ||||| ||||| |||||  
Db 12 AGAGGCGGCGCGCGGCGGTGGAGTTCAGAGA 46

Search completed: August 6, 2002, 21:35:03  
Job time: 6147 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 21:01:12 ; Search time 3320.87 Seconds  
(without alignments)  
12339.085 Million cell updates/sec

Title: US-09-841-963A-1

Perfect score: 1894  
Sequence: 1 gctcgtactctccacgac.....ataaagatactagaactg 1894

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 10010710

Minimum DB seq length: 10  
Maximum DB seq length: 50

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NM\_Main.\*

1:	/cgn2_6/ptodata/2/pna/US085_COMB.seq.*
2:	/cgn2_6/ptodata/2/pna/US06_COMB.seq.*
3:	/cgn2_6/ptodata/2/pna/US07_COMB.seq.*
4:	/cgn2_6/ptodata/2/pna/US080_COMB.seq.*
5:	/cgn2_6/ptodata/2/pna/US081_COMB.seq.*
6:	/cgn2_6/ptodata/2/pna/US082_COMB.seq.*
7:	/cgn2_6/ptodata/2/pna/US083_COMB.seq.*
8:	/cgn2_6/ptodata/2/pna/US084_COMB.seq.*
9:	/cgn2_6/ptodata/2/pna/US085_COMB.seq.*
10:	/cgn2_6/ptodata/2/pna/US086_COMB.seq.*
11:	/cgn2_6/ptodata/2/pna/US087_COMB.seq.*
12:	/cgn2_6/ptodata/2/pna/US088_COMB.seq.*
13:	/cgn2_6/ptodata/2/pna/US089_COMB.seq.*
14:	/cgn2_6/ptodata/2/pna/US090_COMB.seq.*
15:	/cgn2_6/ptodata/2/pna/US091_COMB.seq.*
16:	/cgn2_6/ptodata/2/pna/US092_COMB.seq.*
17:	/cgn2_6/ptodata/2/pna/US093_COMB.seq.*
18:	/cgn2_6/ptodata/2/pna/US094_COMB.seq.*
19:	/cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
20:	/cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
21:	/cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
22:	/cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
23:	/cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
24:	/cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
25:	/cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
26:	/cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
27:	/cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
28:	/cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
29:	/cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
30:	/cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
31:	/cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
32:	/cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
33:	/cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
34:	/cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
35:	/cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
36:	/cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
37:	/cgn2_6/ptodata/2/pna/US100_COMB.seq.*
38:	/cgn2_6/ptodata/2/pna/US101_COMB.seq.*
39:	/cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
40:	/cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
41:	/cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
42:	/cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
43:	/cgn2_6/ptodata/2/pna/US6004_COMB.seq.*

Result No.	Score	Query Match	Length	ID	Description
1	34	1.8	34	US-09-570-593-11	Sequence 11, Appl
2	31	1.6	31	US-09-570-593-10	Sequence 10, Appl
3	24.4	1.3	50	US-09-755-374A-3340	Sequence 3340, Ap
4	24.2	1.3	50	US-09-755-374A-26208	Sequence 26208, A
5	23.6	1.2	30	US-09-755-374A-2825	Sequence 2825, Ap
6	23.6	1.2	48	US-10-017-974-35747	Sequence 35747, A
7	23.2	1.2	47	US-09-422-978-1175	Sequence 1175, App
8	23	1.2	23	US-09-718-159-5	Sequence 5, Appl1
9	23	1.2	23	US-09-718-159-9	Sequence 9, Appl1
10	23	1.2	23	US-09-718-159-10	Sequence 10, Appl
11	23	1.2	23	US-09-718-159-11	Sequence 11, Appl
12	23	1.2	50	US-09-755-374A-9716	Sequence 9716, Ap
13	23	1.2	50	US-06-253-652-21580	Sequence 21580, A
14	22.8	1.2	50	US-09-755-374A-28060	Sequence 28060, A
15	22.6	1.2	38	US-08-472-801-161	Sequence 161, App
16	22.6	1.2	38	US-08-668-235-161	Sequence 161, App
17	22.6	1.2	48	US-09-780-533A-6295	Sequence 6295, Ap
18	22.4	1.2	48	US-09-864-785-3315	Sequence 3315, Ap
19	22.2	1.2	41	US-09-953-198-123	Sequence 123, App
20	22.2	1.2	44	US-08-472-801-999	Sequence 999, App
21	22.2	1.2	44	US-08-668-235-1000	Sequence 1000, Ap
22	22.2	1.2	48	US-10-017-974-22699	Sequence 22699, A
23	22	1.2	22	US-09-064-645-23	Sequence 23, Appl
24	22	1.2	22	US-09-567-970-23	Sequence 23, Appl
25	22	1.2	45	US-09-959-356A-7	Sequence 7, Appl1
26	22	1.2	45	US-09-959-356A-8	Sequence 8, Appl1
27	22	1.2	48	US-09-864-785-3326	Sequence 3326, Ap
28	22	1.2	48	US-10-017-974-34757	Sequence 34757, A
29	22	1.2	50	US-09-755-374A-17016	Sequence 17016, A
30	21.8	1.2	49	US-06-253-652-9804	Sequence 9804, Ap
31	21.8	1.2	50	US-09-755-374A-24550	Sequence 24550, A

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

```
32 21.6 1.1 31 31 US-09-801-274-1269 Sequence 1269, Ap
33 21.6 1.1 47 60 US-60-216-745-4273 Sequence 4273, Ap
34 21.6 1.1 48 35 US-09-930-423-3875 Sequence 3875, Ap
35 21.6 1.1 50 29 US-09-755-374A-23624 Sequence 23624, A
36 21.6 1.1 50 29 US-09-755-374A-23626 Sequence 23626, A
37 21.4 1.1 48 32 US-09-848-754A-7792 Sequence 7792, Ap
38 21.4 1.1 48 33 US-09-864-785-3298 Sequence 3298, Ap
39 21.2 1.1 36 16 US-09-214-913-31 Sequence 31, Appl
40 21.2 1.1 36 23 US-09-612-314-64 Sequence 64, Appl
41 21.2 1.1 36 30 US-09-762-318-11 Sequence 11, Appl
42 21 1.1 21 28 US-09-718-159-3 Sequence 3, Appl
43 21 1.1 37 22 US-09-593-580-45 Sequence 45, Appl
44 21 1.1 37 22 US-09-593-580B-45 Sequence 45, Appl
45 21 1.1 39 30 US-09-771-372-8 Sequence 8, Appl
```

## ALIGNMENTS

```
RESULT 1
US-09-570-593-11/c
; Sequence 11, Application US/09570593
; GENERAL INFORMATION:
; APPLICANT: Kaufmann, Joerg
; APPLICANT: Xin, Hong
; APPLICANT: Hartowe, Greg
; TITLE OF INVENTION: EXPRESSION OF ETS-DOMAIN PROTEINS IN
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 2300-1556
; CURRENT APPLICATION NUMBER: US/09/570,593
; CURRENT FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 60/134,112
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 34
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GSEF-specific primer
US-09-570-593-11
```

```
Query Match 1.8%; Score 34; DB 22; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1850 ccagagatgccccgggaatgataaataaagata 1883
Db 34 CCAGAGATGCCCCGGGAAATGATTAATAAGATA 1
```

```
RESULT 2
US-09-570-593-10
; Sequence 10, Application US/09570593
; GENERAL INFORMATION:
; APPLICANT: Kaufmann, Joerg
; APPLICANT: Xin, Hong
; APPLICANT: Hartowe, Greg
; TITLE OF INVENTION: EXPRESSION OF ETS-DOMAIN PROTEINS IN
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 2300-1556
; CURRENT APPLICATION NUMBER: US/09/570,593
; CURRENT FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 60/134,112
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: GSEF-specific primer
US-09-570-593-10
```

```
Query Match 1.6%; Score 31; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 5e+04;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1552 cagggaggggcaaccaactgccccaggggga 1582
Db 1 cagggaggggcaaccaactgccccaggggga 31
```

```
RESULT 3
US-09-755-374A-3340/c
; Sequence 3340, Application US/09755374A
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide Polymorphisms and M
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: 15966-611
; CURRENT APPLICATION NUMBER: US/09/755,374A
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/174962
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 28742
; SEQ ID NO 3340
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (26)...(0)
; OTHER INFORMATION: 2 of 2 allelic variants (3339 is other entry)
; NAME/KEY: misc_feature
; LOCATION: (25)...(26)
; OTHER INFORMATION: Nucleotide deleted between bases 25 and 26
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Accession number cg44930423
US-09-755-374A-3340
```

```
Query Match 1.3%; Score 24.4; DB 29; Length 50;
Best Local Similarity 68.0%; Pred. No. 9.1e+05;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

```
Oy 399 gcccaaacagcagcgcatgggcagcgccgagccggctgtagcagcga 448
Db 50 GCCTAAAGACAGACAGATAGGCATCCCGAGGCGCTGATGACAGAGA 1
```

```
RESULT 4
US-09-755-374A-26208
; Sequence 26208, Application US/09755374A
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide Polymorphisms and M
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: 15966-611
; CURRENT APPLICATION NUMBER: US/09/755,374A
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/174962
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 28742
; SEQ ID NO 26208
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc-feature
; LOCATION: (26)...(0)
; OTHER INFORMATION: 2 of 2 allelic variants (26207 is other entry)
; NAME/KEY: misc-feature
; LOCATION: (25)...(26)
; OTHER INFORMATION: Nucleotide deleted between bases 25 and 26
; NAME/KEY: misc-feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Accession number cg44009035
US-09-374A-26208
```

```
Query Match          1.3%: Score 24.2; DB 29; Length 50;
Best Local Similarity 71.1%; Pred. No. 9.9e+05;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 968 gccatgctcgaagagcagctccgcacgcgcctcgagtgag 1012
    ||||| || || || || || || || || || || || || || ||
Db 5 ggcacgagggagggcgtgtgtcgcacgagctagcactggaggg 49
```

```
RESULT 5
US-09-755-374A-2825
; Sequence 2825, Application US/09755374A
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide Polymorphisms and Meth
; FILE REFERENCE: 15966-611
; CURRENT APPLICATION NUMBER: US/09/755,374A
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/174962
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 28742
; SEQ ID NO 2825
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (26)...(0)
; OTHER INFORMATION: 1 of 2 allelic variants (2826 is other entry)
; NAME/KEY: misc-feature
; LOCATION: (25)...(26)
; OTHER INFORMATION: Nucleotide deleted between bases 25 and 26
; NAME/KEY: (0)...(0)
; LOCATION: (0)...(0)
; OTHER INFORMATION: Accession number cg43969816
US-09-755-374A-2825
```

```
Query Match          1.3%: Score 23.8; DB 29; Length 50;
Best Local Similarity 72.1%; Pred. No. 1.2e+06;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1124 gacacgaggtgactcatcgtccgcgcgcacccatccacc 1166
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 gccacgctgggtggtcgcacatctcccaagcccccatccacc 47
```

```
RESULT 6
US-10-017-974-35747
; Sequence 35747, Application US/10017974
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Lawrence
; APPLICANT: MCSwigen, James
; TITLE OF INVENTION: Nucleic Acid-Based Treatment of Diseases or Conditions Related to
; FILE REFERENCE: MBH00,1109-A (400/037)
; CURRENT APPLICATION NUMBER: US/10/017,974
; CURRENT FILING DATE: 2001-12-10
```

```
; NUMBER OF SEQ ID NOS: 37080
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35747
; LENGTH: 48
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-10-017-974-35747
```

```
Query Match          1.2%: Score 23.6; DB 37; Length 48;
Best Local Similarity 63.0%; Pred. No. 1.3e+06;
Matches 29; Conservative 3; Mismatches 14; Indels 0; Gaps 0;

Qy 137 ggcctggggtagggagctccctacagcagcagccctgagaccg 182
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ggcugggggaggaacucccucaagcagcaucgucgcggagacggc 46
```

```
RESULT 7
US-09-422-978-175
; Sequence 175, Application US/09422978
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSEP.020cpi
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 175
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-13320-352 : polymorphic base A or G
US-09-422-978-175
```

```
Query Match          1.2%: Score 23.2; DB 18; Length 47;
Best Local Similarity 67.4%; Pred. No. 1.5e+06;
Matches 31; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

Qy 1484 agccagccctgagatggtgggaaacgggcagtgctctgctgt 1529
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 agccaagccctgtaccctcgagargcgcgtgtgtgtctgtgt 46
```

```
RESULT 8
US-09-718-159-5
; Sequence 5, Application US/09718159
; GENERAL INFORMATION:
; APPLICANT: Harkins, Richard
; APPLICANT: Lin, Richard
; APPLICANT: Luke, May
; APPLICANT: Monteciarlo, Felipe
; APPLICANT: Parkes, Deborah
; APPLICANT: Parry, Gordon
; APPLICANT: Steinbrecher, Renate
; APPLICANT: Van Heult, Pamela Toy
; APPLICANT: Xuan, Jian-ai
; TITLE OF INVENTION: DNA Encoding a Novel ETS Polypeptide
; FILE REFERENCE: 51790AUSM1
```

```
; CURRENT APPLICATION NUMBER: US/09/718,159
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/168,182
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: probe
; US-09-718-159-5
```

```
Query Match 1.2%; Score 23; DB 28; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1783 cctgcctccatctgcaccaca 1805
|||||
Db 1 cctgcctccatctgcaccaca 23
```

```
RESULT 9
; Sequence 9, Application US/09718159
; GENERAL INFORMATION:
; APPLICANT: Harkins, Richard
; APPLICANT: Lin, Richard
; APPLICANT: Luke, May
; APPLICANT: Montecarlo, Felipe
; APPLICANT: Parkes, Deborah
; APPLICANT: Parry, Gordon
; APPLICANT: Steindbrecher, Renate
; APPLICANT: Van Heult, Pamela Toy
; APPLICANT: Xuan, Jian-ai
; TITLE OF INVENTION: DNA Encoding a Novel ETS Polypeptide
; FILE REFERENCE: 51790AUSM1
; CURRENT APPLICATION NUMBER: US/09/718,159
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/168,182
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 23
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: antisense
; US-09-718-159-9
```

```
Query Match 1.2%; Score 23; DB 28; Length 23;
Best Local Similarity 73.9%; Pred. No. 1.6e+06;
Matches 17; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1161 tccacctgtgcagttcctcaag 1183
|||||
Db 1 uccaccuuggcagcuuccucaag 23
```

```
RESULT 10
; Sequence 10, Application US/09718159
; GENERAL INFORMATION:
; APPLICANT: Harkins, Richard
; APPLICANT: Lin, Richard
; APPLICANT: Luke, May
; APPLICANT: Montecarlo, Felipe
; APPLICANT: Parkes, Deborah
; APPLICANT: Parry, Gordon
```

```
; APPLICANT: Steindbrecher, Renate
; APPLICANT: Van Heult, Pamela Toy
; APPLICANT: Xuan, Jian-ai
; TITLE OF INVENTION: DNA Encoding a Novel ETS Polypeptide
; FILE REFERENCE: 51790AUSM1
; CURRENT APPLICATION NUMBER: US/09/718,159
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/168,182
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10
; LENGTH: 23
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: antisense
; US-09-718-159-10
```

```
Query Match 1.2%; Score 23; DB 28; Length 23;
Best Local Similarity 91.3%; Pred. No. 1.6e+06;
Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1734 tgactcgacaagccacagca 1756
|||||
Db 1 ugacucgacaagccacagca 23
```

```
RESULT 11
; Sequence 11, Application US/09718159
; GENERAL INFORMATION:
; APPLICANT: Harkins, Richard
; APPLICANT: Lin, Richard
; APPLICANT: Luke, May
; APPLICANT: Montecarlo, Felipe
; APPLICANT: Parkes, Deborah
; APPLICANT: Parry, Gordon
; APPLICANT: Steindbrecher, Renate
; APPLICANT: Van Heult, Pamela Toy
; APPLICANT: Xuan, Jian-ai
; TITLE OF INVENTION: DNA Encoding a Novel ETS Polypeptide
; FILE REFERENCE: 51790AUSM1
; CURRENT APPLICATION NUMBER: US/09/718,159
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/168,182
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 23
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: antisense
; US-09-718-159-11
```

```
Query Match 1.2%; Score 23; DB 28; Length 23;
Best Local Similarity 82.6%; Pred. No. 1.6e+06;
Matches 19; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1699 ccagcattccagagcagcct 1721
|||||
Db 1 ccagcauuuccagagcagccu 23
```

```
RESULT 12
; Sequence 12, Application US/09755374A
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin
```

APPLICANT: Shimkets, Richard A.  
 TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide Polymorphisms and Methods of Use Thereof  
 TITLE OF INVENTION: Use Thereof  
 FILE REFERENCE: 15966-611  
 CURRENT APPLICATION NUMBER: US/09/755,374A  
 CURRENT FILING DATE: 2001-01-08  
 PRIOR APPLICATION NUMBER: 60/174962  
 PRIOR FILING DATE: 2000-01-07  
 NUMBER OF SEQ ID NOS: 28742  
 SEQ ID NO 9716  
 LENGTH: 50  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (26)...(0)  
 OTHER INFORMATION: 2 of 2 allelic variants (9715 is other entry)  
 NAME/KEY: misc\_feature  
 LOCATION: (25)...(26)  
 OTHER INFORMATION: Nucleotide deleted between bases 25 and 26  
 NAME/KEY: misc\_feature  
 LOCATION: (0)...(0)  
 OTHER INFORMATION: Accession number cg43952512  
 US-09-755-374A-9716

Query Match	1.28;	Score 23;	DB 29;	Length 50;
Best Local Similarity	74.4%;	Pred. No. 1.7e+06;		
Matches 29;	Conservative	0;	Mismatches 10;	Indels 0;

Qy 629 agctcgtgcacccaaagcccccctctgtgtgcacgaagtcagag 667  
||||| | | ||| | ||||| ||||| | | |  
Db 1 agctgcacctgcacagaccaaactgtgtgtgcacgaagctgtcag 39

```

RESULT 13
US-60-253-652-21580
: Sequence 21580: Application US/60253652
: GENERAL INFORMATION:
: APPLICANT: Glenn, Matthew
: TITLE OF INVENTION: Compositions isolated from bovine
: FILE OF INVENTION: tissues and methods for their use
: FILE REFERENCE: 1055p2
: CURRENT APPLICATION NUMBER: US/60/253,652
: CURRENT FILING DATE: 2000-11-28
: NUMBER OF SEQ. ID NOS.: 29954
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 21580
: LENGTH: 50
: TYPE: DNA
: ORGANISM: Bovine
US-60-253-652-21580

```

Query Match	1.28;	Score 23;	DB 64;	Length 50;
Best Local Similarity	74.48;	Pred. No. 1.7e+06;		
Matches 29;	Conservative	0;	Mismatches 10;	Indels 0;
			Gaps	0

QY 383 acagacacacgcgccagcccaaacagcagcgcatggc 421  
||| ||||| ||| | ||| | ||| |||||  
Db 12 acctgagacgcgctcacacacatctcggcgccatggc 50

```

RESULT 14
US-09-755-374A-28060
: Sequence: 28060, Application US/09755374A
:
: GENERAL INFORMATION:
:
: APPLICANT: Leach, Martin
:
: APPLICANT: Shumkees, Richard A.
:
: TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide Polymorphisms and Meth
:
: TITLE OF INVENTION: Use Thereof
:
: FILE REFERENCE: 15966-611

```

```

1 CURRENT APPLICATION NUMBER: US/09/755,374A
2 CURRENT FILING DATE: 2001-01-08
3 PRIOR APPLICATION NUMBER: 60/174962
4 PRIOR FILING DATE: 2000-01-07
5 NUMBER OF SEQ ID NOS: 28742
6 SEQ ID NO 28060
7 LENGTH: 50
8 TYPE: DNA
9 ORGANISM: Homo sapiens
10 FEATURE:
11 NAME/KEY: misc_feature
12 LOCATION: (26)..(0)
13 OTHER INFORMATION: 2 of 2 allelic variants (28059 is other entry)
14 NAME/KEY: misc_feature
15 LOCATION: (25)..(26)
16 OTHER INFORMATION: Nucleotide deleted between bases 25 and 26
17 NAME/KEY: misc_feature
18 LOCATION: (0)..(0)
19 OTHER INFORMATION: Accession number cg44009326
20 US-09-755-374A-28060

```

Query Match	1.28;	Score 22.8;	DB 29,	Length 50;
Best Local Similarity	66.08;	Pred. No. 1.8e+06;		
Matches 33; Conservative	0;	Mismatches 17;	Indels 0;	Gaps 0;

**Qy** 485 tcgcgacagccttggagaagcgccagcgggcccagtgtctcgaag 534  
||||| ||||| | ||||| ||||| |||||  
**Db** 1 tcgcgccagcctggcgcgcgaccaggcggcggtgtgttcgttaagtcag 50

```

RESULT 15
US-08-472-801-161/c
: Sequence 161, Application US/08472801
: GENERAL INFORMATION:
: APPLICANT: Hesse 2
: APPLICANT: Smith, Larry J.
: TITLE OF INVENTION: Method and Compositions for Cellular
: TITLE OF INVENTION: Reprogramming
: FILE REFERENCE: Hesse 2
: CURRENT APPLICATION NUMBER: US/08/472,801
: CURRENT FILING DATE: 1995-06-07
: NUMBER OF SEQ ID NOS: 3601
: SOFTWARE: Fastrsq for Windows Version 3.0
: SEQ ID NO 161
: LENGTH: 38
: TYPE: DNA
: ORGANISM: Homo sapiens
US-08-472-801-161

```

Query Match	1.26;	Score 22.6;	DB 8;	Length 38;
Best Local Similarity	75.7%;	Pred. No. 1.9e+06;		
Matches 28; Conservative	0;	Mismatches 9;	Indels 0;	Gaps 0;

QY 958 ggaagtgtgcgcacatgtcggagagacagttccgcag 994  
||||| || ||||| | ||||| ||| |||  
Db 37 GGAGGTCGCCGCAAGCTGGAGGAGCAGGCCCCACAG 1

Search completed: August 6, 2002, 23:36:06  
Job time: 9294 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 21:35:07 ; Search time 378.9 Seconds  
(without alignments)  
10877.386 Million cell updates/sec

Title: US-09-841-963a-1  
Perfect score: 1894  
Sequence: 1 gctgacttcctccacgac.....ataagatcactagagactg 1894

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1362792 seqs, 1088025756 residues  
Total number of hits satisfying chosen parameters: 646124

Minimum DB seq length: 10  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents\_NA\_New:\*  
1: /cgn2\_6/pdata1/pna/PCT\_NEW\_COMB.seq:\*  
2: /cgn2\_6/pdata1/pna/US06\_NEW\_COMB.seq:\*  
3: /cgn2\_6/pdata1/pna/US07\_NEW\_COMB.seq:\*  
4: /cgn2\_6/pdata1/pna/US08\_NEW\_COMB.seq:\*  
5: /cgn2\_6/pdata1/pna/US09\_NEW\_COMB.seq:\*  
6: /cgn2\_6/pdata1/pna/US10\_NEW\_COMB.seq:\*  
7: /cgn2\_6/pdata1/pna/US10\_NEW\_COMB.seq:\*  
8: /cgn2\_6/pdata1/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21.8	1.2	41	US-60-384-980-216	Sequence 216, App
2	21.6	1.1	48	US-09-745-237A-3875	Sequence 3875, Ap
3	21.2	1.1	50	US-09-980-217-51	Sequence 51, Appl
4	21.1	1.1	41	PCT-US02-12063-268	Sequence 268, App
5	21.1	1.1	41	US-10-126-022-268	Sequence 268, App
6	20.8	1.1	48	US-10-156-306-7580	Sequence 7580, Ap
7	20.8	1.1	50	US-09-718-321A-1048	Sequence 1048, Ap
8	20.6	1.1	49	US-09-978-917A-25	Sequence 25, Appl
9	20.6	1.1	49	US-09-997-623-25	Sequence 25, Appl
10	20.4	1.1	48	US-09-745-237A-3883	Sequence 3883, Ap
11	20.2	1.1	41	US-60-384-980-69	Sequence 69, Appl
12	20.2	1.1	50	US-10-099-926-1257	Sequence 1257, Ap
13	20.2	1.1	48	US-09-745-237A-3872	Sequence 3872, Ap
14	20.1	1.1	48	US-09-745-237A-3877	Sequence 3877, Ap
15	19.8	1.0	44	US-09-913-238-45	Sequence 45, Appl
16	19.8	1.0	47	US-10-170-097-1121	Sequence 1121, Ap
17	19.8	1.0	48	US-09-745-237A-4024	Sequence 4024, Ap
18	19.8	1.0	48	US-09-745-237A-4051	Sequence 4051, Ap
19	19.8	1.0	48	US-10-156-306-3945	Sequence 3945, Ap
20	19.8	1.0	48	US-10-156-306-4297	Sequence 4297, Ap
21	19.8	1.0	48	US-10-156-306-7660	Sequence 7660, Ap
22	19.6	1.0	50	US-09-718-321A-1047	Sequence 1047, Ap
23	19.6	1.0	43	US-10-139-480-18	Sequence 18, Appl
24	19.6	1.0	48	US-09-745-237A-3850	Sequence 3850, Ap
25	19.6	1.0	50	US-09-718-321A-228	Sequence 228, App

26	19.4	1.0	40	7	US-10-068-965-5	Sequence 5, Appl1
27	19.4	1.0	41	1	PCT-US02-12063-402	Sequence 402, App
28	19.4	1.0	41	7	US-10-126-022-402	Sequence 402, App
29	19.4	1.0	48	5	US-09-745-237A-4288	Sequence 4288, Ap
30	19.4	1.0	48	7	US-10-156-306-7370	Sequence 7370, Ap
31	19.2	1.0	48	7	US-10-109-349A-102	Sequence 102, App
32	19.2	1.0	48	5	US-09-745-237A-4318	Sequence 4318, Ap
33	19.2	1.0	50	1	PCT-US02-12405-476	Sequence 476, App
34	19.2	1.0	50	7	US-10-144-094-33	Sequence 33, Appl
35	19.2	1.0	50	7	US-10-045-674-476	Sequence 476, Appl
36	19.1	1.0	43	5	US-09-747-774A-80	Sequence 80, Appl
37	19.1	1.0	45	5	US-09-380-447A-246	Sequence 246, Appl
38	19.1	1.0	47	7	US-10-170-097-971	Sequence 971, App
39	19.1	1.0	48	5	US-09-745-237A-3916	Sequence 3916, Ap
40	19.1	1.0	48	5	US-09-745-237A-4015	Sequence 4015, Ap
41	19.1	1.0	48	7	US-10-156-306-4143	Sequence 4143, Ap
42	19.1	1.0	48	7	US-10-156-306-7410	Sequence 7410, Ap
43	19.1	1.0	48	7	US-10-156-306-7653	Sequence 7653, Ap
44	19.1	1.0	50	5	US-09-718-321A-1013	Sequence 1013, Ap
45	18.8	1.0	41	8	US-60-384-980-28	Sequence 28, Appl

## ALIGNMENTS

RESULT 1  
US-60-384-980-216/c  
; Sequence 216, Application US/60384980  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS  
; FILE REFERENCE: D0211 PSP  
; CURRENT APPLICATION NUMBER: US/60/384,980  
; CURRENT FILING DATE: 2002-06-03  
; NUMBER OF SEQ ID NOS: 1219  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 216  
; LENGTH: 41  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-384-980-216

Query Match 1.2%: Score 21.8; DB 8; Length 41;  
Best Local Similarity 70.7%: Pred. No. 3.1e+05;  
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Oy 1233 acaagagagagcattcaaatggagcactgacccag 1273  
Db 41 AAAGGGGAGAGGATGTAAATCTGAGCTCAGGGGACG 1

RESULT 2  
US-09-745-237A-3875  
; Sequence 3875, Application US/09745237A  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Blatt, Larry  
; TITLE OF INVENTION: Met-1  
; FILE REFERENCE: 400/007 (MBH00-918-A)  
; CURRENT APPLICATION NUMBER: US/09/745,237A  
; CURRENT FILING DATE: 2002-04-15  
; NUMBER OF SEQ ID NOS: 4550  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3875  
; LENGTH: 48  
; TYPE: RNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid  
US-09-745-237A-3875

Query Match	1.1%;	Score 21;	DB 1;	Length 41;
Best Local Similarity	82.88;	Pred. No. 4.4e+05;		

Query Match	1.1%	Score 20.8	DB 7	Length 48
Best Local Similarity	60.4%	Pred No. 4	8e+05	
Matches	29	Conservative	2	Mismatches 17
				Indels 0
				Gaps 0
QY	137	gacgtggggaaggaactccctacagacgcgcagccctatagacgcgca	184	
Db	1	ggcagggggggaagaaacucccucacaggaagcaucagucgggggggggca	48	

## RESULT 7

```
US-09-718-321A-1048/c
; Sequence 1048, Application US/09718321A
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: NUCLEIC ACIDS CONTAINING SINGLE NUCLEOTIDE POLYMORPHISMS AND ME
; FILE REFERENCE: 15966-599
; CURRENT APPLICATION NUMBER: US/09/718,321A
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 60/163,783
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 1468
; SOFTWARE: Curagen Patent Formatter Version 0.9
; SEQ ID NO 1048
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: (26)...(0)
; OTHER INFORMATION: single nucleotide polymorphism
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (25)...(26)
; OTHER INFORMATION: Nucleotide deleted between bases 25 and 26
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Accession number cg43955639
US-09-718-321A-1048
```

## Query Match

Best Local Similarity 1.1%; Score 20.8; DB 5; Length 50;  
Matches 31; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 1640 agctgtccctgagagagagagagagagagctgtccccaacacc 1687

DB 49 AGCCGCTACCTGAGAGAGAGAGAGAGAGAGCTGCCAGTACC 2

## RESULT 8

```
US-09-978-917A-25/c
; Sequence 25, Application US/09978917A
; GENERAL INFORMATION:
; APPLICANT: Maxygen Aps; Maxygen Holdings
; TITLE OF INVENTION: Protein C or activated protein C-like molecules
; FILE REFERENCE: 0219us310 - protein C
; CURRENT APPLICATION NUMBER: US/09/978,917A
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 49
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-978-917A-25
```

## Query Match

Best Local Similarity 1.1%; Score 20.6; DB 5; Length 49;  
Matches 29; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 713 gccccagcgagcactgtgctgcccggcgagctgacct 755

DB 46 GCCCCGAGGCGAGCTTGCTGTGTCTCCAGAGACACCT 4

## RESULT 9

## US-09-997-623-25/c

```
; Sequence 25, Application US/09997623
; GENERAL INFORMATION:
; APPLICANT: Maxygen Aps; Maxygen Holdings
; TITLE OF INVENTION: Protein C or activated protein C-like molecules
; FILE REFERENCE: 0219us410 - protein C
; CURRENT APPLICATION NUMBER: US/09/997,623
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US/09/978,917
; PRIOR FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 49
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-997-623-25
```

## Query Match

Best Local Similarity 1.1%; Score 20.6; DB 5; Length 49;  
Matches 29; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 713 gccccagcgagcactgtgctgcccggcgagctgacct 755

DB 46 GCCCCGAGGCGAGCTTGCTGTGTCTCCAGAGACACCT 4

## RESULT 10

```
US-09-745-237A-3833
; Sequence 3833, Application US/09745237A
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease
; FILE REFERENCE: 400/007 (MBH00-918-A)
; CURRENT APPLICATION NUMBER: US/09/745,237A
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 4550
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3833
; LENGTH: 48
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-09-745-237A-3833
```

## Query Match

Best Local Similarity 1.1%; Score 20.4; DB 5; Length 48;  
Matches 28; Conservative 2; Mismatches 16; Indels 0; Gaps 0;

OY 137 ggtgtgggtgagggagcctccctcagcgagcagccctgagacgc 182

DB 1 ggcaggggggaggaacuccuuaagaagacauagcgcggaggggc 46

## RESULT 11

```
US-60-384-980-69/c
; Sequence 69, Application US/60384980
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: D0211 ESP
; CURRENT APPLICATION NUMBER: US/60/384,980
; CURRENT FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 1219
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
```

us-09-841-963a-1.szm.rnp

Search completed: August 6, 2002, 23:42:38

Wed Aug 7 09:11:29 2002

us-09-841-963a-1.szlim.rnpn

Page 5

Job time: 7651 sec

---

